DOCUMENT RESUME

ED 421 892 FL 025 398

AUTHOR Demers, Monique
TITLE La prosodie du discourse rapporte. Recherches sociolinguistiques 3 (The Prosody of Reported Discourse. Sociolinguistic Research 3).
INSTITUTION Laval Univ., Quebec (Quebec). International Center for Research on Language Planning.
ISBN ISBN-2-89219-267-6
PUB DATE 1998-00-00
NOTE 213p.
PUB TYPE Reports - Research (143)
LANGUAGE French
EDRS PRICE MF01/PC09 Plus Postage.
DESCRIPTORS *Discourse Analysis; Foreign Countries; *Intonation; *Language Patterns; Language Research; Language Variation; Oral Language; Regional Dialects; Research Methodology; Social Influences; *Speech Communication; *Suprasegmentals; Syntax
IDENTIFIERS French (Canadian); *Quebec

ABSTRACT

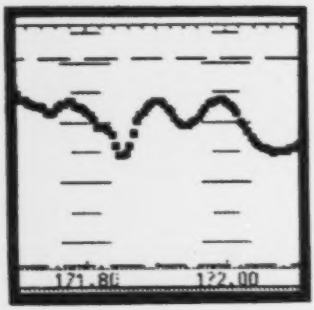
A study investigated the relationship between syntax, discourse, and suprasegmentals in reported utterances (e.g., "He said,..."). The study is intended as both theoretical and descriptive, specifically, proposing a model of intonation in spontaneous oral Quebec French. The first chapter reviews research on Quebec French prosody, on oral spontaneous Quebec French, and on reported discourse. The second chapter presents global principles of intonation and describes the research methodology used in the current study. Chapter three analyzes prosodic production in a corpus of 215 utterances representative of various types of reported discourse. The fourth chapter analyzes perception of the utterances and their prosodic features by 47 different listeners. Contains 121 references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 421 892

MONIQUE DEMERS

LA PROSODIE DU DISCOURS RAPPORTÉ



PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

D. DESHAIES

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

dral

RECHERCHES SOCIOLINGUISTIQUES / 3

FLW2S398

MONIQUE DEMERS

LA PROSODIE DU DISCOURS RAPPORTÉ

CIRAL
Recherches sociolinguistiques/3
Université Laval
Québec
1998

Directrice de la collection
«Recherches sociolinguistiques» : DIANE VINCENT

Éditeur: **CIRAL**
Département de langues et linguistique
Université Laval
Québec (Québec)
G1K 7P4

©MONIQUE DEMERS

ISBN: 2-89219-267-6

Dépôt légal – 1er trimestre 1998
Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

Il est illégal de reproduire cet ouvrage sans autorisation, conformément aux dispositions de la loi sur les droit d'auteur.

TABLE DES MATIÈRES

Table des matières i
Liste des figures vii
Liste des tableaux ix

AVANT-PROPOS 1

INTRODUCTION 3

Chapitre 1
À PROPOS DU DISCOURS RAPPORTÉ

1.1 Le français québécois et l'oral spontané 13
1.1.1 Le français québécois 13
1.1.2 L'oral spontané 15

1.2 Le discours rapporté 16
1.2.1 Le point de vue général de deux grammaires 17
1.2.2 Le point de vue spécifique de quatre études discursives et d'une étude prosodique 19

1.3 Conclusion 22

Chapitre 2
REPRÉSENTATION DE L'INTONATION

2.1 Le choix d'une approche mixte 24
2.1.1 La transcription orthographique 25
2.1.2 La représentation phonétique 26
2.1.3 La représentation phonologique 27

2.2 Le modèle de MERTENS et son application au présent corpus 27
2.2.1 La description formelle des éléments intonatifs 28
Les conventions de notation • Les intervalles mélodiques • Le concept d'accent • La contiguité accentuelle • Les appendices • L'inventaire des formes

2.2.2 Le fonctionnement du regroupement intonatif : la règle de hiérarchie 36

2.3 Conclusion 38

Chapitre 3
ANALYSE DE LA PRODUCTION

3.1 Le corpus 41
3.1.1 La qualité de l'enregistrement 42
3.1.2 Le profil social des informateurs 43
3.1.3 Les types de discours rapporté 44
3.1.4 La délimitation des frontières 46
3.1.5 Le traitement instrumental 47
3.1.6 En somme 49

3.2 Prosodie et syntaxe 50
Les hypothèses • Les résultats
3.2.1 La fréquence fondamentale 52
Le *groupe intonatif avant la transition* • La transition • Le profil global de l'énoncé • L'ampleur des écarts • La cooccurrence des configurations tonales

3.2.2 L'effet des «marqueurs transitionnels» du discours direct sur F_0 62

3.2.3 La pause 69
La fréquence d'occurrence • La durée

3.2.4 Le débit 72

3.2.5 En somme 77

3.3 Prosodie et discours 79
Les hypothèses •
3.3.1 La fréquence fondamentale 85
En discours direct • En discours indirect

3.3.2 La pause 89

3.3.3 Le débit 89
En discours direct • En discours indirect

3.3.4 En somme 93

3.4 Discussion 94
3.4.1 La représentation de l'intonation et l'association prosodie, syntaxe et discours 94

3.4.2 Comparaison avec des discours rapportés tirés du corpus de MERTENS (1987) 97

3.5 Prosodie et facteurs sociaux 99
3.5.1 Quelques études socioprosodiques 99

3.5.2 La distribution des énoncés 104
Les sujets • Les énoncés

3.5.3 Les hypothèses 107

3.5.4 Les résultats 108

3.5.5 Discussion 109
Corrélations observées • Corrélations présentes dans les deux types de discours rapporté *Débit et sexe § Débit et statut socioprofessionnel* • Corrélations présentes dans un seul type de discours rapporté *Intonation et sexe § Intonation et âge § Débit et âge.*

3.6 Conclusion 115

Chapitre 4
ANALYSE DE LA PERCEPTION

PARTIE A : LE CORPUS ORIGINAL

4.1 La méthodologie 121
4.1.1 Le type de test 121
4.1.2 Le protocole 122
4.1.3 Les sujets 122

4.2 Les hypothèses générales 123

4.3 Les résultats
Expérience 1 : Information prosodique et segmentale
4.3.1 Classification syntaxique 125
4.3.2 Classification discursive 129

4.4 Conclusion 130

PARTIE B : UN SOUS-CORPUS

4.5 Le sous-corpus 132
4.5.1 Les énoncés 133
4.5.2 La distinction énoncés dits prosodiquement typiques/ énoncés dits atypiques 133

4.6 La méthodologie 134
4.6.1 Les types de tests 135
4.6.2 La technique de «délexicalisation» 135
4.6.3 Le protocole 137
4.6.4 Les sujets 138

4.7 Les résultats 139
4.7.1 Perception de la différence syntaxique 139
Information prosodique et/ou segmentale non manipulée • *Hypothèses § Expérience 2 § Expérience 3 § Expérience 4 § Expérience 5 § Comparaison § En somme* • Information prosodique manipulée • *Hypothèses § Expérience 6 § Expérience 7 § Expérience 8 § En somme* • Conclusion.
4.7.2 Perception de la différence d'attitude 163
Information segmentale et/ou prosodique • *Hypothèses § Expérience 9 § Expérience 10 § En somme.*

4.8 Conclusion 170

CONCLUSION 173

BIBLIOGRAPHIE 183

ANNEXES 197

10

LISTE DES FIGURES

Chapitre 2

Figure 2.1 Courbe intonative 31

Chapitre 3

Figure 3.1 Courbe intonative 55
Figure 3.2 *Groupe intonatif avant la transition.* Syntaxe .. 56
Figure 3.3 Transition. Syntaxe 57
Figure 3.4 Profil global de l'énoncé. Syntaxe 59
Figure 3.5 Pause silencieuse (LÉANDRI) 68
Figure 3.6 Pause silencieuse. Courbe intonative 68
Figure 3.7 Débit. Vitesse de locution. Syntaxe 74
Figure 3.8 Débit. Vitesse d'articulation. Syntaxe 75
Figure 3.9 *Groupe intonatif avant la transition.* Discours 86
Figure 3.10 Transition. Discours 87
Figure 3.11 Débit. Vitesse d'articulation. Discours 91

Chapitre 4

Figure 4.1 Corpus original. Transition et pause 127
Figure 4.2 Corpus original. Marqueur d'attaque et pause... 128
Figure 4.3 Signal original comparé à un signal filtré, puis «délexicalisé» 136
Figure 4.4 Expérience 2. Sujets naïfs et experts 142
Figure 4.5 Expérience 3. Test-notion et test-transcription 143
Figure 4.6 Expérience 4. Information prosodique ou segmentale 144

Figure 4.7 Expérience 5. Information intégrale 146
Figure 4.8 Pourcentage moyen de perception pour chaque stimulus 147
Figure 4.9 Courbe mélodique originale et modifiée (F_0) 153
Figure 4.10 Expérience 6. Modification de F_0 154
Figure 4.11 Courbe mélodique originale et modifiée (pause) 155
Figure 4.12 Expérience 7. Modification de la pause 156
Figure 4.13 Courbe mélodique originale et modifiée (F_0 et pause) 158
Figure 4.14 Expérience 8. Modification de F_0 et pause 159
Figure 4.15 Expériences 9 et 10 (*accord* ou *désaccord*) 168

i2

LISTE DES TABLEAUX

Chapitre 1

Tableau 1.1 Les relations entre la syntaxe, le discours et la prosodie (GREVISSE:1986;ARRIVÉ *et al.* 1986) 18

Tableau 1.2 Les relations entre la syntaxe, le discours et la prosodie dans les études discursives (AUTHIER:1978; MAINGUENEAU 1981, 1986, 1987; de GAULMYN:1992; VINCENT et DUBOIS 1996) et dans l'étude prosodique (LEANDRI 1993) 19

Chapitre 2

Tableau 2.1 Inventaire des formes dans le modèle de MERTENS 35

Chapitre 3

Tableau 3.1 Composition sociolinguistique du corpus 43

Tableau 3.2 Répartition des discours rapportés dans chaque type de discours et pour chaque locuteur 45

Tableau 3.3 Cooccurrence des domaines significatifs 61

Tableau 3.4 Distribution des «marqueurs transitionnels» 65

Tableau 3.5 Pauses. Fréquence d'occurrence 69

Tableau 3.6 Pauses. Durée 71

Tableau 3.7 Indices de reproduction 81

Tableau 3.8 Distribution des énoncés. Discours 83

Tableau 3.9 Transition. Écarts médians. Discours 88

Tableau 3.10 Débit (DI). Vitesse d'articulation. Discours 92
Tableau 3.11 Configurations tonales. Comparaison avec des discours rapportés tirés du corpus de MERTENS 97
Tableau 3.12 Études socioprosodiques. Tableau-synthèse 102
Tableau 3.13 Distribution des énoncés. Facteurs sociaux..... 106
Tableau 3.14 Corrélations. Facteurs sociaux 110

Chapitre 4
Tableau 4.1 Corpus original. Distribution des sujets. Discours 130
Tableau 4.2 Pourcentage moyen de perception pour chaque situation de perception et chaque catégorie prosodique de discours rapporté 148

AVANT-PROPOS

Observer, comparer, déduire, relancer. Dans la solitude. Écouter, questionner, discuter, argumenter. En compagnie. L'équilibre est possible quand la compagnie est bonne. Et elle l'a été.

L'oreille attentive et les conseils judicieux du phonéticien Claude Paradis, le discours mesuré mais enthousiaste de l'analyste du discours, Diane Vincent, les échanges généreux avec Conrad Ouellon, Denise Deshaies et Jean Dolbec n'ont cessé d'alimenter la présente étude. Merci.

L'apport spécifique des membres de l'Institut de linguistique et de phonétique de Lund en Suède (l'élaboration des tests de perception) et du Centre de recherche de l'Hôpital Saint-Sacrement à Québec (l'analyse statistique) a été déterminant dans la progression de cette recherche. Merci.

Le Conseil de Recherche en Sciences Humaines (CRSH), le Fonds pour la formation de Chercheurs et l'Aide à la Recherche (FCAR, stage à l'étranger), le Fonds de soutien de l'Université Laval, et le projet de recherche PROSO ont appuyé financièrement l'ensemble du travail; l'Université du Québec à Chicoutimi a pour sa part contribué à la publication. De tels programmes, petits ou grands, apportent non seulement un soutien financier nécessaire, mais aussi une motivation profonde. Merci.

INTRODUCTION

On parle beaucoup. On connaît pourtant bien peu les mécanismes de cette activité complexe qu'est la parole. On veut les comprendre davantage ? Il faut alors se limiter à un tout petit rouage, mais pour qu'il soit utile, il faut pouvoir placer le petit rouage dans le grand système. Le défi est lancé.

Par rapport à la syntaxe et à la sémantique, la phonétique constitue un champ d'étude à la limite du linguistique. Et lorsque le linguiste s'approche de la phonétique, c'est en principe pour examiner quels sont les rapports que la phonétique entretient avec le «coeur syntaxico-sémantique» du système linguistique. C'est encore plus vrai si le linguiste s'intéresse à des éléments prosodiques comme la mélodie ou le débit. Il sait que les systèmes de synthèse de la parole ont du mal à se passer de la prosodie. Il en déduit qu'il peut difficilement considérer la prosodie comme une redondance de la syntaxe. Pourtant, un énoncé monotone n'est pas non plus tout à fait incompréhensible. Il se met alors à penser que le va-et-vient entre les niveaux syntaxique, discursif et prosodique est constant. Reste à comprendre comment ils se complètent... Parfois ces niveaux travaillent en duo, ou en trio, parfois chacun travaille en solo...

Dans la présente étude, le matériau pour l'observation des rapports entre syntaxe, discours et prosodie a d'abord été choisi en fonction de la structure syntaxique; celle-ci devait pouvoir être clairement définie tout en présentant des constructions variées. Ensuite, cette structure syntaxique devait pouvoir être porteuse de valeurs discursives variées. Finalement, la richesse prosodique devait être à tout le moins «pressentie». Le discours rapporté répond à ces conditions.

D'une part, tous les grammairiens reconnaissent l'indépendance syntaxique comme caractéristique du discours direct et la subordination comme caractéristique du discours indirect. La définition morpho-syntaxique du discours rapporté est donc claire et l'identification des formes directe et indirecte pose peu de problèmes. En même temps, le discours rapporté directement présente régulièrement des formes syntaxiques variées, tantôt les paroles rapportées sont introduites par un verbe *dire*, comme dans l'exemple (1) :

(1) Elle a beaucoup pleuré pi elle a dit : «C'est dommage les enfants qui t'aimaient tant.» (6227-79)[1]

tantôt elle sont introduites sans verbe *dire*, comme dans l'exemple (2):

(2) Là c'est rendu : «Maman tu fais rien on s'en va au restaurant maman.» (4850ø-59)

ou encore, les paroles citées débutent par des éléments lexicaux comme *ah, bon, ben*, etc., comme dans l'exemple (3) :

(3) J'ai pas de poste que je dis : «**Ah** non je veux pas écouter ça.» (6136-131)

Est-ce qu'il y aura une relation entre ces diverses constructions syntaxiques et les caractéristiques prosodiques ?

D'autre part, des utilisations variées de la structure morpho-syntaxique du discours rapporté, particulièrement celle du discours direct, donnent à penser que ces utilisations représentent des types

---

[1] Le nombre **gras**, entre parenthèses, indique qu'il s'agit d'un exemple tiré du corpus de l'étude (Montréal 84). La lettre **grasse** entre parenthèses indiquera que les exemples sont tirés d'un autre corpus que celui ici étudié. L'énoncé est identifié de la façon suivante : numéro de l'énoncé, indice du type de discours rapporté (l'absence d'indice indique qu'il s'agit d'un discours direct standard, c'est-à-dire introduit par le verbe *dire*; [ø] = discours direct sans verbe *dire*; [i] = discours indirect) et le numéro du locuteur. Les conventions de transcription orthographique sont présentés dans la section 2.1.1.

d'emploi discursif variés. Les exemples (4) à (8) ont en commun la structure morpho-syntaxique du discours direct, mais d'un point de vue discursif, ils sont différents.

(4) Il m'a dit : «Ça été l'erreur de ta vie.» (5188-73)
(5) Je me disais : «Mon-Dieu je suis dans mon bout'.» (4663-59)
(6) C'est gênant se faire dire : «Vous avez pris du poids.» (4043-15)
(7) Pour moi vous allez dire : «À soir elle les hommes là.» (4034-15)
(8) C'est pour ça 'je dis moi : «Il-y-a pas de femme heureuse comme moi.» (5553-108)

L'exemple (4) correspond au discours rapporté «classique». Le locuteur premier n'est pas le rapporteur et les paroles réfèrent à une situation d'énonciation antérieure (*a dit*). À partir de l'exemple (5), les paroles rapportées débordent les cadres de cette définition. Dans l'exemple (5), le rapporteur et l'énonciateur se confondent. L'exemple (6) suggère plutôt que les paroles rapportées constituent une sorte de vérité répétable dans plusieurs situations. L'exemple (7) représente des paroles qui sont anticipées alors que les paroles rapportées de l'exemple (8) pourraient tout aussi bien être introduites par *il semble que* que par le verbe *dire*. La forme du discours rapporté apparaît donc porteuse de valeurs discursives variées. Est-ce qu'il y aura une relation entre ces valeurs discursives variées et les caractéristiques prosodiques ?

Ces observations conduisent à une définition élargie du discours rapporté. Est ici considéré comme discours rapporté tout acte d'énonciation, introduit ou non par une verbe de parole (*dire*), qui rapporte, directement ou indirectement, des propos tenus par un autre locuteur, ou par l'énonciateur lui-même, que ces propos aient été prononcés réellement ou imaginés.

Le marquage prosodique d'un discours rapporté (par exemple, par l'imitation de la voix, de l'intonation, du débit du locuteur cité), particulièrement d'un discours rapporté directement, a été produit, ou perçu, un jour ou l'autre par n'importe quel locuteur. Certains analystes du discours rapporté (dont AUTHIER:1978; MAINGUENEAU:1981) font état, plus ou moins intuitivement, de cette place qu'occupe la prosodie dans le discours rapporté. D'autres, comme de GAULMYN (1992), sont plus précis quant aux conditions d'apparition des marques prosodiques. À ma connaissance, une seule étude quantitative et instrumentale de la prosodie du discours rapporté a cependant été effectuée (LÉANDRI:1993). L'étude porte sur du français hexagonal lu et met en évidence les différences prosodiques liées à la structure syntaxique du discours direct par rapport à celle du discours indirect. Les différences prosodiques entre les types de discours rapporté sont bel et bien existantes, mais les connaissances demeurent fragmentaires. Dans quelle mesure les caractéristiques prosodiques du discours rapporté sont fonction de la construction syntaxique et/ou de la valeur discursive de l'énoncé ?

Toutes ces caractéristiques (une définition syntaxique claire, des constructions syntaxiques variées, des valeurs discursives variées et des marques prosodiques observées) font du discours rapporté une matière prometteuse pour l'étude des relations entre les différents niveaux linguistiques.

Après avoir choisi d'étudier la structure du discours rapporté, les critères généraux de mise en place du corpus se sont imposé d'emblée : du français québécois, oral spontané, produit par des locuteurs représentatifs des différents groupes sociaux.

La prosodie du français québécois est encore mal connue. Les premiers travaux marquants (GENDRON:1966; BOUDREAULT:1968, 1970) remontent aux années 60. Il s'agissait alors d'études faites à partir de phrases lues, par des locuteurs cultivés. Peu de chose par la suite. Il a fallu attendre la fin des années 80 avant que des équipes (celles de Santerre, de Cedergren, de Paradis) n'entreprennent des

recherches intensives sur la prosodie du français québécois. Les corpus étudiés se sont diversifiés. De plus en plus d'études sont effectuées à partir de l'oral spontané produit par des locuteurs représentatifs des différents groupes sociaux. La présente étude s'inscrit dans ce courant et devrait apporter une modeste contribution aux connaissances sur le fonctionnement de la prosodie du français québécois. Plus spécifiquement, une étude discursive du discours rapporté en français québécois oral spontané a récemment été effectuée par VINCENT et DUBOIS (1996). Il est apparu intéressant de reprendre une partie du corpus utilisé pour en faire l'étude prosodique.

Par ailleurs, un certain nombre d'études non spécifiques au discours rapporté (dont MEHTA et CUTLER:1988; GUAÏTELLA:1991; BLAUUW:1994) montrent que, règle générale, la prosodie est plus variée, plus «riche» en spontané qu'en lu. Il est raisonnable de penser que les caractéristiques prosodiques du discours rapporté oral spontané seront plus variées que celles qui ont été observées en lu.

Finalement, la relation entre des caractéristiques prosodiques et des facteurs sociaux a déjà été démontrée par plusieurs études (dont SACHS:1975; RYALLS *et al.*:1994; BYRD:1994). Il a semblé nécessaire de vérifier dans quelle mesure les tendances prosodiques du discours rapporté pouvaient être dépendantes, ou indépendantes, de facteurs sociaux tels le sexe, l'âge et le statut socioprofessionnel.

La présente étude prosodique du discours rapporté est instrumentale. Elle prend en compte la fréquence fondamentale (F_0), la pause silencieuse et le débit. Cependant, la façon d'extraire les valeurs de F_0 constitue à elle seule tout un débat. La façon de représenter l'intonation ne fait pas non plus consensus. Les différentes propositions de représentation de l'intonation sont donc examinées de manière à adopter la représentation la plus appropriée au type de corpus ici traité. L'analyse prosodique s'appuie sur une classification syntaxique des discours rapportés puis sur une classification discursive. Cette analyse prend également en compte des facteurs extra-

linguistiques. Les caractéristiques prosodiques observées sont ensuite vérifiées par des tests de perception.

La question générale qui se pose maintenant est la suivante : la prosodie du discours rapporté en français québécois oral spontané agit-elle en conjonction avec des facteurs linguistiques tels la structure syntaxique, et/ou une valeur discursive, et/ou certains facteurs extra-linguistiques ?

OBJECTIF

L'objectif général poursuivi est à la fois théorique et descriptif. Il est théorique parce que l'étude porte sur les relations entre la prosodie, la syntaxe et le discours. Mais l'objectif est aussi, et principalement, descriptif parce que l'analyse vise d'abord à décrire et à caractériser la prosodie de l'énoncé rapporté. La prosodie de ce dernier est en premier lieu considérée au niveau phonétique ou acoustique : le signal produit est analysé indépendamment de son rôle linguistique. Par la suite, la prosodie du discours rapporté est bien sûr considérée au niveau phonologique : le signal produit est analysé en fonction de son rôle linguistique sur les plans syntaxique et discursif, en fonction d'éventuelles influences de facteurs sociaux puis en fonction de la corrélation entre le signal produit et le signal perçu.

La réalisation d'un tel objectif comporte une importante difficulté méthodologique en ce qu'elle requiert la mise au point d'une méthode jusqu'à un certain point «mécaniste» pour traiter un matériau éminemment fluide, la parole spontanée. En bref, cette méthode est établie de la façon suivante :

1) Examen des modèles prosodiques qui ont déjà été appliqués à l'étude du français : celui de GARDING, de HIRST et DI CRISTO ainsi que de MERTENS.

2) Choix d'une approche mixte, plutôt rigide (règles fixes pour la prise de mesure de F_0, calcul des écarts en demi-tons,

application de seuils de perception), afin d'uniformiser un matériau très diversifié (diversité dans la longueur des énoncés, dans la construction syntaxique, diversité également au plan des locuteurs ainsi que des situations de communication).

3) Vérification de la validité de la méthodologie retenue pour l'analyse de la production par une analyse de la perception.

4) Vérification du degré de signification des résultats par une analyse statistique.

QUESTIONS SPÉCIFIQUES

Du précédent objectif découlent les questions spécifiques suivantes :

1) Quel est le modèle de représentation de l'intonation le plus approprié au traitement du présent corpus de français québécois oral spontané ?

En parole spontanée :

2) Est-ce qu'il y a une distinction entre la structure syntaxique qu'est le discours direct et la structure syntaxique qu'est le discours indirect du point de vue de la fréquence fondamentale, de la pause et du débit ?

3) En discours direct, la structure prosodique est-elle sensible à la variation syntaxique propre à l'énoncé rapporté en parole spontanée : présence ou absence du verbe introducteur (*dire*) ainsi que présence ou absence d'un «marqueur transitionnel»[2] ?

---

[2] Le terme «marqueurs transitionnels» désigne ici les pauses et les marqueurs d'attaque tels *ah, ben, coudon,* etc. que l'on peut retrouver à la transition entre le discours citant et le discours cité en discours direct.

4) Les caractéristiques prosodiques sont-elles fonction de la valeur discursive de l'énoncé rapporté ? Si oui, jusqu'où la finesse prosodique peut-elle aller dans le marquage des valeurs discursives ? Est-ce qu'il y a un marquage binaire, expressif ou neutre, ou un marquage n+aire, rattaché à une valeur de reproduction, plus ou moins narrative, ou encore rattaché à des valeurs «attitudinales» multiples ?

5) Dans quelle mesure les traits prosodiques produits sont-ils fonction de facteurs sociaux tels le sexe, l'âge et le statut socioprofessionnel ?

6) Dans quelle mesure les caractéristiques prosodiques produites sont-elles perçues ?

SOMMAIRE DES CHAPITRES

Le **chapitre 1** présente brièvement l'état des connaissances sur la prosodie du français québécois, sur les particularités prosodiques de l'oral spontané ainsi que sur les caractéristiques syntaxiques, discursives et prosodiques du discours rapporté. Dans le cas du français québécois et de l'oral spontané, les faits relevés ne concernent pas le discours rapporté, mais ils ont une incidence sur le sujet traité.

En ce qui concerne spécifiquement le discours rapporté, le point de vue de grammaires, d'études discursives et d'une étude prosodique est examiné de manière à faire ressortir les relations qui ont déjà été décrites entre la syntaxe, la valeur discursive et les caractéristiques prosodiques du discours rapporté.

Dans le **chapitre 2**, il y a présentation globale des principaux modèles de représentation de l'intonation ainsi que de l'approche mixte retenue. L'objectif descriptif de l'étude (la description des caractéristiques prosodiques du discours rapporté) et les moyens instrumentaux disponibles (le logiciel de traitement du signal acoustique

Computerized Speech Lab) ont fait opter pour une représentation dans laquelle la notation à l'oreille de MERTENS (1987) est remplacée par les valeurs phonétiques. La façon d'extraire F_0 est empruntée à ROSSI (1971 et 1978), les écarts intra- et intersyllabiques sont calculés en demi-tons ('t HART *et al.*: 1990) alors que la représentation abstraite (ton haut, ton bas), telle qu'elle est proposée par MERTENS (1987), est conservée. En raison de son application à la présente étude, l'ensemble des éléments du modèle de MERTENS est présenté de façon détaillée : conventions de notation, intervalles mélodiques, concept d'accent et de découpage intonatif, contiguïté accentuelle, appendices, inventaires des formes et règle de hiérarchie tonale.

Le **chapitre 3** porte sur l'analyse de la production. Le corpus étudié est issu d'un corpus d'entrevues sociolinguistiques, le corpus de Montréal 1984 (THIBAULT et VINCENT: 1990). Il est constitué de 215 énoncés représentatifs des types de discours rapporté : le discours direct avec verbe *dire*, le discours direct sans verbe *dire* et le discours indirect. Ces énoncés sont produits par 16 locuteurs représentatifs de trois facteurs sociaux : l'âge, le sexe et le statut socioprofessionnel. Pour l'âge et le statut socioprofessionnel, les valeurs extrêmes ont été retenues. La répartition des discours directs et des discours indirects ainsi que la délimitation des énoncés s'appuient sur des critères syntaxiques, sémantiques et discursifs.

L'analyse de la production prosodique est effectuée pour les trois paramètres suivants : la fréquence fondamentale, la pause et le débit. En ce qui concerne l'analyse de la fréquence fondamentale, il a été décidé de considérer plusieurs domaines : le discours citant, le *groupe intonatif avant la transition*, la transition, le discours cité, le *groupe intonatif après la transition* et le profil global de l'énoncé. L'effet des «marqueurs transitionnels» sur F_0 est aussi observé. Les pauses sont examinées en termes d'occurrence et de durée. Quant au débit, il est calculé de deux façons: en considérant la vitesse de locution (incluant les temps de pause) puis la vitesse d'articulation (excluant les temps de pause).

Cette analyse met successivement en relation la prosodie et la syntaxe, la prosodie et le discours, la prosodie et les facteurs sociaux. Toutes les analyses sont appuyées sur des analyses statistiques appropriées.

Le **chapitre 4** traite de l'analyse de la perception et se divise en deux parties. Dans la première partie, le corpus original (215 énoncés) est soumis à une quarantaine d'auditeurs. Il s'agit de mettre en relation le pourcentage d'auditeurs qui «reconnaîtront» le type de discours rapporté, en discours direct ou en discours indirect, et les caractéristiques prosodiques de fréquence fondamentale, de pause et de débit. Dans la deuxième partie, un petit nombre d'énoncés est manipulé. Une dizaine d'énoncés représentatifs de discours prosodiquement typiques et atypiques sont présentés sous diverses formes : une mélodie sans mots, des mots sans mélodie puis des mélodies modifiées (F_0 et pause silencieuse à la transition). Au total, 9 expériences sont réalisées. Sept d'entre elles portent sur la distinction grammaticale du discours direct par rapport au discours indirect. Parmi celles-ci, quatre sont effectuées à partir de l'information prosodique et/ou segmentale non modifiée alors que les trois autres sont effectuées à partir d'une information prosodique modifiée. Finalement, deux expériences sont réservées à la distinction d'une attitude, celle de l'*accord* ou du *désaccord* du rapporteur avec les paroles citées. Toutes les analyses sont fondées sur des analyses statistiques.

On retrouve en annexe un exemple des explications et des consignes données lors d'une expérience de perception des différences grammaticales (*Annexe A*) ainsi qu'un exemple des explications et des consignes données lors d'une expérience de perception des différences d'attitude (*Annexe B*). L'enregistrement du corpus ainsi que la transcription orthographique, phonétique et phonologique de l'entier du corpus sont disponibles sur demande.

Chapitre 1

À PROPOS DU DISCOURS RAPPORTÉ

Le discours rapporté a été, et demeure, l'objet de nombreuses études. Il ne s'agit pas d'en faire une revue exhaustive et détaillée, mais plutôt de faire ressortir les relations entre les caractéristiques syntaxiques, discursives et prosodiques qui ont orienté mon étude. Mais auparavant, quelques remarques générales sur la prosodie du français québécois et de l'oral spontané sont requises en raison de l'incidence de ces deux composantes sur le sujet traité : la prosodie du discours rapporté, en français québécois, oral spontané.

1.1 LE FRANÇAIS QUÉBÉCOIS ET L'ORAL SPONTANÉ

1.1.1 LE FRANÇAIS QUÉBÉCOIS

À la fin des années 60, la question fondamentale des phonéticiens d'ici demeure : le *français du Canada* est-il prosodiquement différent du *français de France*[1] ? Au cours des années 70, on parlera plutôt de français québécois, peut-être plus d'ailleurs pour des raisons idéologiques que pour des raisons linguistiques. Quoi qu'il en soit, à partir de ce moment, la question fondamentale devient plutôt : quelles sont les caractéristiques prosodiques du français québécois ?

[1] On évitera l'appellation *français standard* pour désigner le français de l'Hexagone : il n'est pas de standard que le français de France.

Dans le même courant, la présente étude est orientée de façon à montrer les caractéristiques du discours rapporté en français québécois plutôt qu'à montrer l'importance de la variété dialectale dans la prosodie du discours rapporté.

Bien que nous ne disposions, à ce jour, que de résultats épars et parfois contradictoires concernant des caractéristiques de la prosodie québécoise, certaines d'entre elles ne peuvent être passées sous silence, en raison de leur possible incidence sur le sujet traité.

D'une part, d'un point de vue intonatif, l'état des connaissances sur la présumée monotonie du registre (étendue et hauteur de la voix) français québécois[2] demande tout de même une certaine prudence quant à l'utilisation des seuils de perception des écarts intersyllabiques, seuils issus d'une étude réalisée à partir d'un corpus de français hexagonal (MERTENS 1993), la position ici adoptée est discutée dans la section 2.2.1).

D'autre part, d'un point de vue socioprosodique, quelques études québécoises, qui ne portent pas, bien sûr, sur le discours rapporté (il n'en existe pas), attirent tout de même l'attention sur le fait que le schéma *classique* de l'énoncé déclaratif (montant-descendant) est moins fréquent en oral spontané qu'en lu (BALIGAND et JAMES:1973). Ces observations sont intéressantes en regard du discours rapporté. Est-ce que le *classique* schéma montant-descendant de la phrase énonciative, apparié par les grammairiens à celui du

---

[2] Les études de VINAY (1955), CHEMERIS (1966), BOUDREAULT (1968, 1970), effectuées à partir de corpus lus, par des locuteurs cultivés, affirment que le registre des Québécois est moins étendu que celui des Français. BOUDREAULT (1968:116) a toutefois signalé que les différences observées pouvaient être reliées au «niveau culturel» ou au sexe. À partir de spontané radiophonique, produit par des locuteurs français et par des locuteurs franco-ontariens, tous cultivés, HOLTZER (1968) n'observe pas cette différence de registre. Plus récemment, à partir de l'étude de bulletins de nouvelles, BISSONNETTE (1995:48) observe plutôt que les lecteurs québécois ont un registre plus étendu que les lecteurs français.

discours indirect, se maintiendra en spontané ? Est-ce que le *classique* schéma d'intonation montante observé par LEANDRI (1993) en discours direct se maintiendra en spontané ? D'autres analyses ont montré que l'effet de contexte, lu ou spontané, peut être différent selon qu'il s'agit d'hommes ou de femmes (MAURY et WRENN 1973), ou encore que le statut professionnel influence l'étendue des variations intonatives (HOLDER 1968 191; LEON et MARTIN 1970 82). Mais sur ce dernier point, les études se contredisent. Malgré tout cela, il convient de se demander : est-ce que les variations d'intonation du discours rapporté refléteront ces distinctions liées au sexe ou à l'appartenance sociale ?

1.1.2 L'ORAL SPONTANÉ

Pourquoi la présente étude prosodique porte-t-elle sur de l'oral spontané ? Parce que de façon générale, c'est la forme orale la plus courante, toutefois, c'est la forme orale la moins connue, particulièrement du point de vue prosodique. En effet, la plupart des études prosodiques portent sur des corpus lus; GUAITELLA (1991 65) cite à peine une trentaine d'études réalisées à partir du spontané. Des théories générales de l'intonation sont même fondées sur des corpus lus, pourtant peu représentatifs de la réalité, comme le déplorent LIEBERMAN *et al.* :

> The data further suggest that read speech may not be the most appropriate material on which to base or test theories for the linguistic functions of intonation. [...] Reading is a comparatively recent invention in human history, moreover, most human communication is spontaneous, not read from text. (LIEBERMAN, KATZ, JONGMAN, ZIMMERMAN et MILLER 1985 656)

En ce qui concerne le discours rapporté, on verra que l'intonation du discours rapporté a également surtout été décrite à partir d'exemples fabriqués, voire littéraires, que la seule étude prosodique effectuée à ce

jour porte sur un corpus lu[3]

Dans diverses études, citées par MEHTA et CUTLER (1988:136), il apparaît que les caractéristiques prosodiques de la parole spontanée sont distinctes de celles de la parole lue. De plus, la variation prosodique est plus importante en spontané qu'en lu (BLAUUW:1995). Il faut par ailleurs être conscient qu'il y a différents types de lu et différents types de spontané. En contexte lu, il y a certainement des différences entre un texte lu en laboratoire, la lecture d'une histoire à un enfant et la lecture d'un texte spontané déjà produit par le lecteur; de même, en contexte spontané, la confidence risque de se distinguer de l'entrevue télévisée. Il reste cependant que, de façon générale, le patron prosodique fixe, ou régulier, est plus associé à la lecture qu'au spontané.

1.2 LE DISCOURS RAPPORTÉ[4]

La structure du discours rapporté est étudiée en vue d'analyser les rapports qu'entretiennent la syntaxe, le discours et la prosodie. Il importe donc de faire la synthèse de ce que l'on sait du discours rapporté à ces trois niveaux linguistiques. D'une part, les grammaires[5], même si elles s'intéressent davantage à l'écrit et présentent de ce fait une vision

---

[3] Il est à noter qu'une étude sur l'intonation du discours rapporté direct dans l'oral spontané (MOREL: 1996) vient de paraître. Il en sera fait brièvement état dans la section 4.7.2.

[4] Le discours indirect libre a été exclu de la présente étude. Les bases syntaxiques de mon étude (le DIL ne présente pas de marques syntaxiques univoques) et le contexte d'oral spontané (selon MAINGUENEAU (1986:95), c'est dans la littérature romanesque qu'il est utilisé «au maximum de ses possibilités») s'y prêtaient peu.

[5] Le point de vue de la grammaire générative n'est pas ici traité. Une étude détaillée du discours rapporté a été effectuée dans ce cadre théorique (FORGET 1980).

moins adaptée au corpus d'oral spontané ici traité, permettent pourtant d'établir des distinctions de base entre le discours direct et le discours indirect. Des analyses discursives approfondies permettent d'autre part de préciser les liens entre syntaxe, discours et prosodie. Au plan strictement prosodique, il n'existe, à ma connaissance, qu'une seule étude quantitative du discours rapporté (LEANDRE 1993).

1.2.1 LE POINT DE VUE GÉNÉRAL DE DEUX GRAMMAIRES

Parmi les grammaires consultées, certaines (dont celles de WAGNER et PINCHON 1962, de CHARAUDEAU 1992, de LE GOFFIC 1993) présentent des connaissances relativement développées sur le discours rapporté du point de vue énonciatif, mais ne mentionnent d'aucune façon l'aspect prosodique alors que la grammaire de GREVISSE (1986) et celle d'ARRIVÉ, GADET et GALMICHE (1986) font état, même si c'est de façon sommaire, de la dimension prosodique de l'énoncé rapporté. Comme il s'agit de la question centrale de la présente étude, le discours des grammairiens sur le discours rapporté est présenté à travers *Le bon usage* et *La grammaire d'aujourd'hui*.

En premier lieu, il ressort que la définition du discours rapporté que donnent ces grammaires est plus étroite que celle que j'utilise. Alors que celles-ci désignent comme du discours rapporté uniquement des paroles prononcées dans une situation d'énonciation antérieure, j'inclus dans la notion de discours rapporté tout ce qui utilise la forme syntaxique du discours rapporté.

Le *Tableau 1.1* résume les relations entre la syntaxe, le discours et la prosodie du discours rapporté telles que présentées par ces deux grammaires.

Tableau 1.1 Les relations entre la syntaxe, le discours et la prosodie du discours rapporté telles que présentées dans GREVISSE (1986) et dans ARRIVÉ *et al.* (1986).

	DD	DI
Syntaxe	Deux indépendantes	Une complétive
	⇓	⇓
Discours	Absence du point de vue du narrateur	Présence du point de vue du narrateur
	⇓	⇓
Prosodie	Particulière (ARRIVÉ *et al.*)	Fixe

Dans la grammaire de GREVISSE comme dans la grammaire d'ARRIVÉ *et al.*, il apparaît que la structure syntaxique influence la valeur discursive, qui elle-même influence la prosodie : le discours direct est constitué de deux indépendantes, ce qui entraîne l'absence de point de vue du narrateur et une prosodie particulière pour les paroles rapportées; le discours indirect constitue pour sa part une complétive, ce qui correspond à une présence du point de vue du narrateur et une prosodie fixe, celle d'un énoncé déclaratif. Syntaxe, discours et prosodie vont de pair.

Cependant, l'inclusion d'énoncés qui ont la forme syntaxique du discours rapporté sans correspondre à des paroles prononcées antérieurement rompt la correspondance entre la syntaxe et la valeur discursive. Dans le discours direct de l'exemple (1), on voit tout de suite que le narrateur n'exclut pas son point de vue, bien au contraire.

(1) Mais ce que je dis *je dis* : «On: on chôme pour faire travailler (pause) les Japonais les Coréens (pause) pi les Chinois» (4144-30)

Le contrecoup sur la prosodie est probable. Il est légitime de se demander si l'«intonation particulière» attribuée au discours cité directement sera maintenue dans des paroles citées sur-le-champ, comme dans cet exemple (1) ?

1.2.3 LE POINT DE VUE SPÉCIFIQUE DE QUATRE ÉTUDES DISCURSIVES ET D'UNE ÉTUDE PROSODIQUE

Le *Tableau 1.2* permet d'illustrer les précisions apportées par les études discursives et l'étude prosodique tout en mettant en parallèle les divers types de relations observées entre syntaxe, discours et prosodie.

Tableau 1.2 Les relations entre la syntaxe, le discours et la prosodie du discours rapporté telles que présentées dans les études discursives (A, B et C) et dans l'étude prosodique (D).

(A) Études discursives : AUTHIER (1978) et MAINGUENEAU (1981, 1986, 1987)

CORPUS ÉCRITS	DD	DI
Syntaxe	Deux indépendantes	Une complétive
Discours	Mise à distance	Prise en charge
	↔	
	Manifestations de l'attitude du rapporteur dans les deux types de discours rapporté.	
	↓	↓
Prosodie	Variable (toutes les nuances, mais perception aléatoire) (AUTHIER)	Peu utilisée (le lexique est souvent le marqueur d'attitude)

(B) Étude discursive : de GAULMYN (1992)

CORPUS ORAL SPONTANÉ	DD	DI
Syntaxe	Deux indépendantes	Une complétive
	↔	
Discours	Forme expressive ou forme neutre	
	↓	
Prosodie	Souvent remplacée par des éléments lexicaux (*oh, ah, bon, ben*, etc.)	
	↔	
	La courbe mélodique est souvent identique	

(C) Étude discursive : VINCENT et DUBOIS (1996)

CORPUS ORAL SPONTANÉ	DD	DI
Syntaxe	Deux indépendantes	Une complétive
	↔	
Discours	Valeur de reproduction, de *(+) narrative* à *(-) narrative* (les critères de narrativité sont reliés au temps verbal, à l'identification des interlocuteurs et de la situation)	
Prosodie	Pas de mention	

(D) Étude prosodique : LÉANDRI (1993)

CORPUS LI	DD	DI
Syntaxe	Deux indépendantes	Une complétive
	↓	↓
Prosodie	Pause	
	Pente montante	Pente descendante

La relation entre syntaxe et discours

Dans les études discursives, la forme syntaxique du discours rapporté n'est pas absolument liée à la valeur discursive. Le discours indirect, aussi bien que le discours direct, peut être porteur de diverses valeurs discursives. Celles-ci ne font cependant pas l'unanimité. Les valeurs attribuées sont reliées à :

- une valeur d'attitude, quasi variable à l'infini (AUTHIER et MAINGUENEAU);
- une valeur expressive ou une valeur neutre (de GAULMYN);
- une valeur de reproduction décomposable en cinq types d'emploi (VINCENT et DUBOIS).

La relation entre syntaxe et prosodie

Dans l'étude prosodique, la valeur discursive n'est pas prise en compte. La forme syntaxique du discours rapporté est liée à une valeur prosodique (LÉANDRI). Les caractéristiques prosodiques du discours direct sont distinctives de celles du discours indirect.

La relation entre discours et prosodie

L'association entre la valeur discursive et les caractéristiques prosodiques est traitée différemment selon les études. AUTHIER et MAINGUENEAU accordent une grande importance à la prosodie : elle pourrait véhiculer une infinité d'attitudes («de l'agressivité polémique à la complicité tendre» (AUTHIER 1978:70). Il faut ajouter que les auteurs tirent leurs observations de corpus écrits et que AUTHIER précise que la perception de la prosodie est «aléatoire». De GAULMYN limite pour sa part la correspondance entre le discours et la prosodie : ou bien le discours rapporté est expressif, ou bien il est neutre; selon l'auteure, en spontané, il arrive souvent que les marques morphologiques ou mimo-gestuelles remplacent les marques prosodiques. VINCENT et DUBOIS ne prennent pas en compte les valeurs prosodiques.

Dans mon étude, le lien entre le discursif et le prosodique sera examiné selon l'idée de continuum proposée par VINCENT et DUBOIS. L'hypothèse de départ est que plus le discours rapporté correspondra à une valeur narrative forte (temps verbal au passé, interlocuteurs identifiés et situation précisée), plus il sera susceptible de porter des caractéristiques prosodiques marquées et inversement. L'adoption de cette valeur discursive distribuée selon cinq types d'emploi constitue un point milieu entre la valeur discursive référant à de nombreuses attitudes non répertoriées (AUTHIER:1978; MAINGUENEAU:1981, 1986, 1987) et à celle référant à une valeur discursive binaire, expressive ou neutre (de GAULMYN:1992). La classification des cinq types d'emploi ne va pas sans quelques problèmes (cf. section 3.3, exemple (13), mais elle constitue, à ce moment-ci, la façon la plus rigoureuse de tenir compte de la valeur discursive.

1.3 CONCLUSION

La revue de quelques études sur le français québécois, l'oral spontané et le discours rapporté a permis de préciser les aspects de la question étudiée. L'étude est orientée de manière à montrer les caractéristiques du discours rapporté en français québécois plutôt qu'à montrer l'importance de la variété dialectale dans la prosodie du discours rapporté. La présumée différence de registre entre la variété québécoise et la variété hexagonale est cependant traitée avec attention en raison de l'utilisation de seuils de perception issus d'une étude de français hexagonal. Il ressort des études sur l'oral spontané, qui ont été examinées, que la prosodie du spontané est moins stéréotypée, ou en tout cas moins prévisible, que celle de la lecture oralisée.

Du point de vue central des liens éventuels entre la syntaxe, le discours et la prosodie en discours rapporté, la revue de grammaires, d'études discursives et d'une étude prosodique a permis de tracer les pistes de recherche. Il apparaît que la valeur discursive de l'énoncé rapporté n'est pas forcément liée à la structure syntaxique du discours rapporté d'une part et que cette valeur discursive peut avoir une influence sur la prosodie, influence à tout le moins aussi importante que celle exercée par la syntaxe, d'autre part.

Chapitre 2

REPRÉSENTATION DE L'INTONATION

Quel type de représentation de l'intonation utiliser pour analyser l'organisation prosodique des énoncés ? C'est un problème épineux. Il y a absence remarquable de consensus. Les principales propositions qui ont actuellement cours peuvent être subdivisées en deux grands types d'approche.

La première, de type prédictif, s'inscrit dans la continuité de la grammaire générative et propose des modèles soit à partir de la structure syntaxique (comme PIERREHUMBERT 1980), soit à partir de la structure métrique (comme DELL 1984). La proposition de PIERREHUMBERT demeure sans doute la plus utilisée. L'auteure soutient que le seul niveau phonétique de représentation est celui de la courbe de F_0 elle-même, elle rejette toute représentation stylisée, interpolée, de la courbe de F_0, contrairement aux propositions descriptives présentées ci-après. Le système de transcription utilisé est ToBI (Tones and Break Indices). Bien que d'une grande précision, ce système a été conçu pour l'anglais et ne peut s'appliquer tel quel au français[1]. Ce type d'approche n'est pas retenu pour la présente étude.

---

[1] L'équipe de Cedergren, de l'Université du Québec à Montréal, travaille actuellement à l'adaptation de ToBI pour le français.

La seconde approche, de type descriptif, prend son départ aux faits phonétiques pour rendre compte des aspects physiologiques, acoustiques et perceptuel avant d'en faire une interprétation proprement linguistique, c'est-à-dire syntaxique et énonciative, puis de construire des règles prédictives. Dans ce type d'approche, qui paraît plus adapté pour l'étude de l'oral spontané, au moins trois grandes propositions peuvent être utilisées. La prise de mesure de F_0 peut être effectuée à l'aide d'instruments et porter sur les voyelles (ROSSI:1971 et 1978), ou sur des points-cibles (GARDING:1977, 't HART, COLLIER et COHEN:1990, HIRST et DI CRISTO (à paraître), ou encore, l'intonation peut également être mesurée de façon impressionniste, à l'oreille (MERTENS :1987).

Les modèles de GARDING, de HIRST et DI CRISTO et celui de MERTENS ont été expressément appliqués au français. Après examen de ces modèles (cf. DEMERS 1996:59-68), le choix effectué pour la présente étude se porte sur une approche qualifiée de mixte dont le modèle de MERTENS constitue le noyau.

2.1 LE CHOIX D'UNE APPROCHE MIXTE

Compte tenu de l'objectif, à savoir l'analyse de l'organisation prosodique du discours rapporté en français québécois oral spontané, du point de vue de la production, et compte tenu de l'instrumentation disponible, il a été décidé d'opter pour une approche mixte. La représentation phonétique est exprimée par les valeurs acoustiques (contrairement au modèle de MERTENS) alors que la représentation phonologique est faite à l'aide de quatre tons (comme dans le modèle de MERTENS). Les groupes intonatifs délimités par un accent final se regroupent en unités plus grandes selon une règle de hiérarchie tonale (MERTENS:1987, cf. 2.2.2). Dans ce qui suit, chaque type de représentation est présenté : la transcription orthographique, la représentation phonétique et le représentation phonologique.

2.1.1 LA TRANSCRIPTION ORTHOGRAPHIQUE

Ce premier niveau est commun à toutes les propositions de représentation de l'intonation. Il n'existe cependant pas de conventions universelles reconnues pour la transcription d'un corpus oral. Les principes ici retenus pour la transcription orthographique respectent la plupart des conventions utilisées dans THIBAULT et VINCENT (1990). Le but de cette transcription est de faciliterle repérage des formes et des structures. Le principe de transcription suivant est retenu : la transcription se conforme généralement aux règles de l'orthographe standard.

Par exemple, du point de vue phonétique, (i) il n'est pas tenu compte des phénomènes phonétiques propres à l'oral tels la réduction de deux syllabes à une seule syllabe (c'était), (ii) de la disparition d'un phonème (ils), (iii) de la prononciation particulière d'une unité ([aʒetra] est transcrit achètera). Par ailleurs, les ajouts sont notés : (iv) la prononciation non standard de la consonne finale est notée par une apostrophe détachée à droite (bout'), (v) une liaison non étymologique est marquée par la présence de la consonne détachée (ça va_t être). (vi) En raison de leur prononciation quasi uniforme par l'ensemble des locuteurs, les particules discursives telles *puis, bien, tu sais* sont transcrites *pi, ben, tsé*; les sacres *christ, calice, hostie, tabernacle* sont transcrits *criss, câlice, osti, tabarnak*.

Du point de vue syntaxique, (i) les morphèmes qui sont phonétiquement absents de l'énoncé sont marqués par une apostrophe détachée à gauche et à droite ('faut' je me trouve une job); (ii) le phonème «disparu» qui laisse des traces acoustiques est transcrit (il le dit)[2]; (iii) la transformation phonétique usuelle d'une construction syntaxique est transcrite entre parenthèses (je lui (y) dis); (iv) les homophones de *tu* et de *plus* sont distingués de la façon suivante : *tu*

---

[2] Il est à signaler que l'absence du *ne* n'a pas été marquée par l'apostrophe, le repérage des phrases négatives pouvant être fait par les particules *pas* ou *plus*.

pour le pronom personnel et *tU* pour le morphème interrogatif (tu viens-tU) ; *plus* pour le marqueur additif ou comparatif et *plUs* pour le morphème de négation; (v) les items lexicaux qui prennent valeur d'expression sont marqués par des traits d'union (Fait-que); (vi) les troncations, les faux-départs, les reprises de syllabes de même que les tenues sont marqués par le signe phonétique de l'allongement [:] (La par: quand on donne sa parole); (vii) le début et la fin de la partie citée du discours direct sont marqués par les [:] et les guillemets [« »]. Le point [.] ou le point d'interrogation [?] sont utilisés pour marquer la fin des énoncés. Aucun autre signe de ponctuation n'est utilisé; (viii) enfin, les rires sont notés entre parenthèses et les chevrons délimitent les phénomènes de *back-channel* (signal d'écoute).

2.1.2 LA REPRÉSENTATION PHONÉTIQUE

La représentation phonétique a été établie à l'aide des valeurs mesurées instrumentalement, comme dans toutes les propositions de représentation de l'intonation, excepté celle de MERTENS (1987). La prise de mesure se fait selon les conventions suivantes : (i) relevé des valeurs de durée syllabique et pausale (ms); (ii) relevé des valeurs de F_0 (Hz) aux deux tiers de la durée vocalique, ce qui a pour principal avantage d'éviter la majeure partie des effets microprosodiques co-intrinsèques; (iii) calcul des écarts fréquenciels en demi-tons (DT) avec des seuils de perception. Les seuils de perception des intervalles intersyllabiques (ou plus précisément intervocaliques puisque, dans ce cas-ci, la mesure est prise à l'intérieur de la voyelle) mineurs et majeurs sont principalement établis à partir des seuils proposés par MERTENS (1987). Ce transfert en DT permet de comparer les comportements prosodiques des locuteurs dont les gammes tonales sont différentes alors que la prise en compte de seuils de perception permet de s'approcher des éléments linguistiquement pertinents.

Les inconvénients liés au présent type de représentation phonétique, perte des paramètres prosodiques autres que la fréquence et la pause, ont été compensés par une révision perceptuelle faite par

deux experts. En bout de ligne, la transcription finale devrait être assez proche, autant que faire se peut, de ce qui est produit et perçu.

2.1.3 LA REPRÉSENTATION PHONOLOGIQUE

La représentation phonologique, ou abstraite, de l'énoncé suit le modèle de MERTENS, notation à partir de 4 niveaux de hauteur : haut, bas, suraigu et infra-bas. Ce modèle offre l'immense avantage d'être très explicite: une structure claire du groupe intonatif, des critères formels de distinction entre l'accent initial et l'accent final (même s'ils ne s'appliquent pas de façon absolue dans le présent corpus) ainsi que d'identification de l'appendice, des règles de contiguité accentuelle et, surtout, une règle de hiérarchie des tons finals permettant le regroupement des groupes intonatifs en unités sémantiques plus larges. La justification du choix de ce type de représentation phonologique sera plus profitable au fur et à mesure de la présentation du modèle de MERTENS, ce qui fait l'objet de la section suivante.

2.2 LE MODÈLE DE MERTENS ET SON APPLICATION AU PRÉSENT CORPUS

Dans l'ensemble, le modèle proposé par MERTENS (1987) est construit à partir des mêmes grands principes que ceux de GARDING et de HIRST et DI CRISTO :

1) Marques temporelles et fréquencielles pour ce que MERTENS appelle l'*accent final* (AF) correspondant à l'*accent de groupe* (AG), non final et final dans le modèle de GARDING, à l'*accent rythmique primaire* (R_1) et à l'*accent nucléaire* (R_0) dans le modèle de HIRST et DI CRISTO,
2) seule marque fréquencielle pour ce qu'il désigne comme *accent initial* (AI), correspondant à l'*accent focal* (AF) dans le modèle de GARDING et à l'*accent rythmique secondaire* (R_2) dans le modèle de HIRST et DI CRISTO.

3) En revanche, le modèle de MERTENS propose un regroupement des groupes intonatifs, ce que ne fait pas le modèle de GARDING. Dans le modèle de HIRST et DI CRISTO, le regroupement intonatif fait référence aux valeurs temporelles et tonales pour le regroupement rythmique (l'*accent rythmique primaire* et l'*accent nucléaire*, hiérarchisés selon leur valeur temporelle et tonale respective) et aux valeurs tonales pour le regroupement tonal (*accent «rythmique» secondaire*). Dans le modèle de MERTENS, le regroupement intonatif s'effectue à partir de la seule valeur tonale des *accents finals*, selon une règle systématique de hiérarchie tonale.

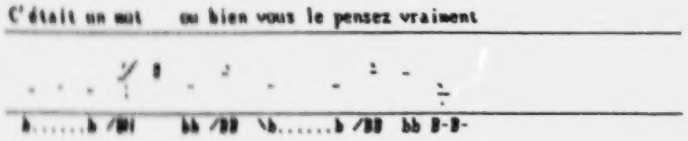
Il convient maintenant de présenter plus en détail les éléments formels de l'intonation tels que décrits par MERTENS (1987:69-122) et l'application qui en a été faite dans la présente étude.

2.2.1 LA DESCRIPTION FORMELLE DES ÉLÉMENTS INTONATIFS

Les conventions de notation

Les conventions de notation proposées par MERTENS se divisent en trois parties : une transcription orthographique, une transcription auditive et une transcription abstraite.

(a)



(MERTENS 1990:160)

Dans la présente étude, la transcription orthographique, comme la transcription abstraite, est parenthétisée selon la délimitation des

groupes intonatifs afin de faciliter la perception visuelle de la structure prosodique par rapport à la structure syntaxique. En ce qui concerne la transcription auditive, elle a été remplacée par des valeurs acoustiques : les valeurs temporelles (ms) et fréquencielles (Hz et DT). La représentation phonologique se fait à l'aide de quatre niveaux de hauteur: le niveau haut (H ou h), bas (B ou b), suraigu (H+) et infra-bas (B- ou b-); la double majuscule (ex. : HH) représente l'accent final, la simple majuscule (ex. : H), l'accent initial et la minuscule (ex. : h ou hh), la syllabe non accentuée.

L'exemple (1) illustre les trois niveaux de représentation, orthographique, phonétique et phonologique (TP signifie transcription prosodique), tels que transcrits dans la présente approche mixte.

	(Orthographique)										
(1)	(Il m'a dit) :			(«Ç'a été)			(l'e rreur)			(de ta vie.») (5188-73)	
	(Phonétique)										
ms	54	129	204	145	162		136	283		142 192 195	
Hz	89	84	116	79	79		75	85		70 67 68	
DT			5,59	-6,65				2,16		-3,36	
	(Phonologique)										
TP	(b		b HH)	(bb	BB)	(bb		/BB)	(	B-B-)	

Les intervalles mélodiques

Dans le système de MERTENS, l'intervalle majeur, c'est-à-dire le changement de niveau (de bas à haut ou vice-versa), requiert «typiquement» 5 DT (MERTENS 1993 23) alors que l'intervalle mineur requiert entre 2 et 3 DT (MERTENS 1993 23). C'et intervalle mineur est indiqué comme un rehaussement [/] ou un abaissement [\]. L'auteur précise toutefois que l'intervalle varie en fonction du locuteur et en fonction de facteurs phonostylistiques.

Étant donné l'état actuel des connaissances sur les différences éventuelles de registre entre le français québécois et le français hexagonal (cf. section 1.1.1), doit-on utiliser tels quels les seuils proposés par MERTENS, seuils issus d'un corpus de français hexagonal?

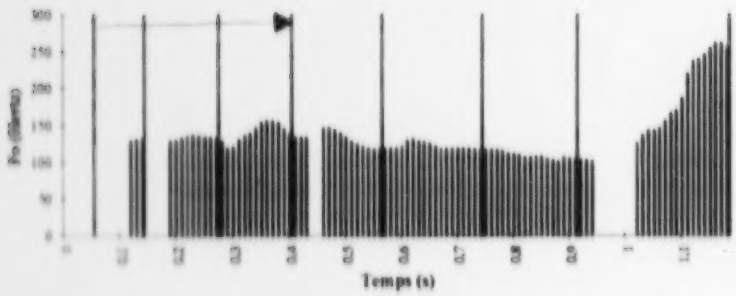
Les positions des études québécoises sont diverses[1]. Dans le cas présent, c'est à partir des données empiriques que le seuil d'intervalle mineur a été fixé à 2 DT et le seuil d'intervalle majeur à 4 DT. Évidemment, des études avec un plus grand nombre de locuteurs, et un plus grand nombre d'énoncés, sont requises afin de déterminer si le seuil de perception varie en fonction du dialecte, ou du contexte, ou du registre du locuteur... Quant à la hauteur de la première syllabe d'un énoncé, elle est notée en fonction de la fréquence moyenne du locuteur.

Les tons dynamiques sont relevés sur l'étendue de la voyelle selon les seuils proposés par ROSSI (1971 et 1978) et DI CRISTO (1985), l'écart intervocalique est calculé à partir de la mesure prise aux deux tiers, comme celui de la voyelle statique, la transcription est faite selon les conventions suivantes : BH ou HB quand l'intervalle est majeur, soit supérieur ou égal à 4 DT, et B/B, B\B ou H/H, H\H quand l'intervalle est mineur, soit entre 2 et 3 DT.

La *Figure 2.1* présente la courbe intonative d'un énoncé accompagnée de la représentation de l'intonation selon les conventions présentées ci-haut. On remarquera dans cet exemple un accent final /HH et un autre BH. Dans le premier cas, il s'agit d'un ton statique (pas de variation intravocalique perceptible) et d'un intervalle intervocalique mineur (2,25 DT), dans le second cas, il s'agit d'un ton dynamique et d'un intervalle intervocalique majeur (15,01 DT).

---

[1] DEBAILLY (1994) choisit d'abaisser l'intervalle mineur à 1,5 DT et l'intervalle majeur à 3 DT. Ce choix suppose que le registre en français québécois (FQ) soit moins étendu qu'en français de France (FF). À partir de la perception de 20 énoncés d'oral spontané produits par un Québécois et de 20 énoncés produits par un Français (les deux locuteurs sont de sexe masculin, sont dans la trentaine et appartiennent à la classe moyenne), DESHAIES et PARADIS (1995) fixent pour leur part le seuil de perception des syllabes proéminentes (ce qui correspond à l'intervalle mineur) à 3DT pour le FQ et à 4 DT pour le FF. Ces résultats tendraient également à démontrer que le registre du FQ serait moins étendu que celui du FF.



2)	**(J'ai**	**pas**	**dit)**	**(que j'en**	**pren**	**drai**	**pas.)**	
ms		129	127	165	177	174	266	
Hz	130	137	156	122	118	100	*(122-257)*238	[4]
DT			2,25	-4,26		-2.87	*(12,9)*15,01	
TP	(h.....h		/HH)	(b.....b		\bb	BH)	

Figure 2.1 Courbe intonative et représentation de l'intonation. Courbe intonative de *J'ai pas dit que j'en prendrai pas* (6621i-99).

Le concept d'*accent* et le découpage intonatif

MERTENS distingue deux types de syllabes accentuées : celles porteuses d'un *accent final* (AF) et celles porteuses d'un *accent initial* (AI), terminologie reliée aux positions accentuables, la fin et le début d'un constituant[5], les autres syllabes sont désignées comme non

---

[4] Le changement de ton à l'intérieur de la voyelle (ton dynamique) est noté de la façon suivante : le début et la fin des valeurs fréquencielles vocaliques sont en italiques, entre parenthèses; la valeur aux deux tiers du ton dynamique est notée après la parenthèse fermante, sans attribut. L'écart en DT en italiques, entre parenthèses, indique l'écart intravocalique alors que l'écart qui suit la parenthèse fermante, sans attribut, indique l'écart intersyllabique.

[5] Pour la distinction entre les multiples désignations de ces deux catégories d'accent, on se reportera avec profit à l'article de Rossi (1979:17-18).

accentuées (NA). L'*accent final*[6] a une fonction démarcative, c'est-à-dire qu'il marque la fin d'un groupe intonatif, il est bien sûr obligatoire. Il porte les caractéristiques prosodiques suivantes : il accepte les tons statiques (haut, bas, suraigu et infra-bas), les tons dynamiques, ou glissandi (montant et descendant) et l'allongement; il peut également être suivi, mais non précédé, d'une pause silencieuse.

L'*accent initial*[7], à valeur informative ou expressive, facultatif, se place généralement à l'initiale du mot. Toutefois, comme l'avait déjà observé MAROUZEAU (1924), MERTENS signale que ce n'est pas toujours le cas : «Quand la syllabe commence par une voyelle, l'*accent initial* passe parfois à la deuxième syllabe [...] Il y a même quelques cas isolés de mots à initiale consonantique, accentués sur la deuxième syllabe [...]» (MERTENS 1992:150). PASDELOUP (1990) observe pour sa part, dans un corpus de phrases lues, des cas d'*accent initial*[8] sur l'antépénultième. Selon elle, environ 20% des cas d'accent d'insistance, ou d'accent *initial*, se présentent ailleurs que sur la syllabe initiale. Pour DI CRISTO (à paraître), la distribution de l'*accent initial* est reliée à des contraintes phonotactiques, principalement au nombre de syllabes dans le groupe intonatif. Quoi qu'il en soit, selon MERTENS, l'*accent initial* se caractérise prosodiquement de la façon suivante : un ton statique haut, rarement bas, et bref (noyau vocalique, l'attaque

---

6 Dans la tradition britannique, cet accent à valeur démarcative est désigné comme un *accent nucléaire* (nuclear stress, cf. CRYSTAL 1991). DI CRISTO (à paraître) comme CRUTTENDEN (1986) utilisent, dans ce cas, la dénomination d'*accent primaire*.

7 Cet accent à valeur informative ou expressive connaît de multiples désignations, entre autres : *accent focal*, au sens de focus étroit (CRUTTENDEN 1986; COUPER-KUHLEN 1986; BRUCE et TOUATI 1990) ou *accent secondaire* (DI CRISTO, à paraître), *accent externe* (ROSSI 1985).

8 PASDELOUP (1990), comme HIRST et DI CRISTO (à paraître), utilise plutôt l'appellatif *accent secondaire*.

consonantique est souvent importante, contrairement à l'*accent final*), il peut être précédé, mais non suivi d'une pause silencieuse.

Si l'ensemble des accents initiaux du présent corpus se conforment à la description prosodique qu'en fait MERTENS, deux cas d'*accent initial* sont toutefois porteurs de noyaux vocaliques longs par rapport aux noyaux vocaliques adjacents et de tons dynamiques (DEMERS 1996:74-78).

En ce qui concerne la brièveté de l'*accent initial*, la preuve de l'absence d'allongement comme caractéristique absolue reste encore à donner. On peut cependant penser que la durée joue un rôle beaucoup moins important dans la caractérisation de l'*accent initial* que dans celle de l'*accent final*. D'autre part, même si la présence du ton dynamique n'est peut-être pas tout à fait étrangère à cette «longueur», la caractérisation absolue de l'*accent initial* par le ton statique ne peut pas non plus être ici maintenue, caractéristique par ailleurs non utilisée ni par DI CRISTO (à paraître), ni par TORIATI (1987).

Pour MERTENS, «Une syllabe est accentuée si elle sonne avec plus de relief que les syllabes environnantes. Il s'agit donc ici d'une proéminence par le niveau sonore ou par la force phonatoire, [...]» (MERTENS 1993:24). Il est par ailleurs précisé que le renforcement d'intensité devient redondant dans les cas où l'*accent final* est porteur de caractéristiques qui lui sont réservées (l'allongement, les tons dynamiques et le ton suraigu). Dans le corpus à l'étude, l'ensemble de ces traits perceptuels est remplacé par les seuils de perception et par la durée syllabique en millisecondes pour l'allongement d'une syllabe par rapport à celles qui précèdent. Ce premier regroupement établi à partir de la production est ensuite révisé à l'écoute par deux experts. Si une syllabe est porteuse de l'indice mélodique alors que la suivante est porteuse de l'indice temporel, la transcription représente la perception des juges, s'ils sont en accord; autrement, elle est faite selon l'indice mélodique.

La contiguïté accentuelle

La distinction *accent initial/accent final* permet d'expliciter les cas de contiguïté accentuelle, exclus par DELL (1984) et PASDELOUP (1988). La typologie des syllabes accentuées contiguës est abondamment expliquée par MERTENS (1992)[9]. Les séquences possibles, avec ou sans pause silencieuse, sont :

AF + (#) + AF (2 groupes intonatifs)
AI + (#) + AF (1 groupe intonatif)
AF + (#)+ AI (2 groupes intonatifs)

Dans le corpus ici analysé, tous ces cas de figure ont été rencontrés, sauf AI + # + AF (également non illustré dans le corpus de MERTENS), la séquence la plus fréquente est de loin AF + AF.

Les appendices

> MERTENS définit l'appendice de la façon suivante :
> [une syllabe ou] une série de syllabes non accentuées (...) non suivie d'une syllabe accentuée (...). Ce type de série atone, appelée «appendice», a certaines caractéristiques propres. D'abord sa forme plane : on n'y trouve ni de montée ni de descente. Ensuite, son niveau de hauteur, limité au ton **b-** [en note, «et **h** dans d'autres cas»]. Enfin l'appendice ne peut apparaître qu'après les tons **B-B-**, **HB-**, **HB**, **BH** et **H+H+**. Les syllabes atones précédant l'accent ne connaissent pas ces contraintes. (MERTENS 1990:166).

Le présent corpus en contient une dizaine dont tous sont des éléments optionnels, mot ou hésitation.

---

[9] SANTERRE (1990) a également observé le phénomène de contiguïté accentuelle en français québécois. Cependant, l'observation de SANTERRE ne recouvre que les accents de nature différente : accent de nature mélodique (AI) + accent de nature durative (AF).

L'inventaire des formes

La structure du groupe intonatif correspond à la suivante : seul l'*accent final* (AF) est obligatoire; les syllabes non accentuées (NA) —incluant l'appendice— ainsi que l'*accent initial* (AI) sont facultatifs:

(NA) AI) (NA) AF (NA)(appendice)
(MERTENS 1990:167)

Les formes attribuables à chaque catégorie accentuelle constitutive du groupe intonatif sont données dans le *Tableau 2.1*.

Tableau 2.1 Inventaire des formes dans le modèle de MERTENS (1990:169)

((NA)	AI)	(NA)	AF	(appendice))
			B-B-	b-b-
			H+H+	
			HB-	
			HB	
b	H	b		
h	B	h	BH	b-b-
			\HH	
				hh
			HH	
			BB	

Cette description formelle des éléments intonatifs est complétée par une interprétation sémantique du système intonatif. Sur une base empirique, une hiérarchie des tons a été établie. En effet dans le système proposé par MERTENS, le type d'*accent final*, variable selon l'intention du locuteur, permet de regrouper ou non un groupe intonatif avec le groupe intonatif qui le suit et d'établir ainsi des liens sémantiques à partir uniquement des valeurs tonales. C'est ce que présente la section suivante.

2.2.2 LE FONCTIONNEMENT DU REGROUPEMENT INTONATIF : LA RÈGLE DE HIÉRARCHIE

On reconnaît généralement le rôle organisateur de l'intonation au niveau de l'énoncé, bien que les explications sur le fonctionnement du mécanisme diffèrent considérablement (cf. DELL 1984; MARTIN 1981). Le dépouillement du corpus d'oral spontané a permis à MERTENS (1987) d'établir des règles d'organisation intonative basées uniquement sur la hiérarchie des tons à l'*accent final*. Cette organisation est purement intonative puisqu'elle peut fonctionner indépendamment de la structure syntaxique; elle est systématique parce qu'elle s'applique aussi bien aux phrases ambiguës, aux phrases à structure syntaxique brisée, fréquentes en spontané, qu'aux phrases classiquement grammaticales. La hiérarchie des tons est la suivante :

1. L'infra-bas est au sommet de la hiérarchie. Le suraigu domine les tons hauts et bas. Quant au ton haut, il domine le ton bas.

   B-B- > H+H+ > HH > BB

2. Les tons dynamiques se rangent entre l'infra-bas et les tons statiques, mais c'est la partie dominante à l'intérieur de la syllabe qui détermine la valeur hiérarchique du ton entier.

   HB- = B-B-
   HB = BH = HH

3. Le ton rehaussé domine le ton sans rehaussement correspondant; le ton sans abaissement domine le ton abaissé correspondant.

   /HH > HH
   HH > \HH

L'application de la hiérarchie des tons se fait de gauche à droite:

a) Quand le ton de l'accent final du premier groupe est inférieur à celui du deuxième groupe, il y a regroupement des deux groupes intonatifs;

b) quand le premier ton de l'accent final est supérieur à celui du deuxième, ou quand les deux tons sont identiques, il y juxtaposition, c'est-à-dire non-inclusion des deux groupes.

L'exemple (3) permettra d'illustrer le fonctionnement de l'organisation intonative selon la règle de hiérarchie. Les parenthèses ( ) indiquent les groupes intonatifs alors que les crochets [ ] marquent les regroupements sémantiques. Dans l'exemple (3), le premier groupe est identique au deuxième, il y a donc juxtaposition. Le deuxième groupe est inférieur au troisième (B < H), il y a donc inclusion; le troisième groupe est lui-même inférieur au quatrième (BH < B-B-), il y a encore inclusion.

(3)

[(Elle me les passe)][(elle me les emmène)(«Veux-tu les lire?»)(Mais &[mb 660 ms][10]]
4074ø-15

TP [(...../BB)][(..... /BB)(..... BH)(..... B-B-)]
1 2 3 4

Il est à remarquer que des éléments qui font partie intégrante de l'oral, tels les faux-départs, les troncations, les ponctuants, les pauses d'hésitation, souvent mis de côté dans les études prosodiques, trouvent ici leur place : leur valeur prosodique est notée puis soumise à la hiérarchie des tons. Ils sont ainsi rattachés sémantiquement à un groupe intonatif. La troncation *chp* de l'exemple (4), bien qu'elle s'insère dans la partie citante du discours rapporté, fait partie de l'unité sémantique contenant le début du discours cité, parce que sa valeur mélodique est inférieure à celle du groupe intonatif suivant.

(4)

[(Pi dans)][#161 ms][11](moi &[chp 130 ms])(«Non)] [(il est tout à fait)] [(normal »)]
5712ø-121

Hz 128 105 151
TP [...../BB)] [(.....BB)(HH)] [(...../BB)] [(...../BB)]

---

[10] La perluète [&] correspond à une pause d'hésitation, on trouve en exposant l'émission sonore ainsi que sa durée.

[11] Le dièse [#] correspond à une pause silencieuse, le fait qu'il soit en exposant indique que la durée du silence n'est pas perceptible (< 250 ms).

La pause d'hésitation peut parfois être accentuée et former un groupe intonatif autonome. (MERTENS, 1993:44), comme en (5).

(5)

[(Dès qu'ils voulaient aller)(dans une direction) # <hum>(: &$^{euh\ 518\ ms}$)][(Va z y.»)]
5189₀-73

Hz		150	*(146-88)*130	84	149	
TP	[(..... BB)(..... HH)	(H\H)][(.....HH)]				

2.3 CONCLUSION

La représentation de l'intonation ne faisant pas l'unanimité parmi les chercheurs, diverses approches ont été globalement examinées puis une approche mixte pour la représentation phonétique (valeurs acoustiques) et phonologique (modèle de MERTENS:1987) du présent corpus a été retenue.

Le système de MERTENS est issu de l'analyse d'un corpus de français oral spontané. De façon générale, les règles d'application de son modèle sont les plus précises, donc les plus facilement applicables, en même temps que les plus adéquates à une description complète du corpus à l'étude. Par exemple, les troncations, répétitions, hésitations y sont non seulement notées, mais aussi hiérarchisées selon leur valeur tonale. Ces «éléments embarrassants» du discours spontané peuvent alors être rattachés à des unités sémantiques grâce à des critères exclusivement prosodiques. L'analyse prosodique d'une structure aussi marquée syntaxiquement que celle du discours rapporté, à travers du discours spontané, requiert un maximum de systématicité dans la représentation de l'intonation. Le modèle de MERTENS, avec quelques adaptations, est apparu comme le plus adéquat.

Le remplacement de la transcription auditive, tel que faite dans le modèle de MERTENS, par les valeurs acoustiques révisées à l'écoute devrait conférer à la représentation finale une représentation assez

proche de ce qui est produit en même temps que de ce qui est perçu. De plus la règle de hiérarchie permet le regroupement sémantique des groupes intonatifs sur la base de leur valeur tonale. En même temps que ce modèle met en évidence une certaine correspondance entre la structure syntaxique et la structure intonative, il illustre la liberté du locuteur d'ajouter des accents aux frontières syntaxiques mineures ou d'utiliser plus ou moins l'*accent initial* par exemple. Un jeu largement exploité en oral spontané.

Chapitre 3

ANALYSE DE LA PRODUCTION

Avant de passer à l'analyse de la production prosodique proprement dite, les composantes du corpus à l'étude sont présentées : la qualité de l'enregistrement, le profil social des informateurs, les types de discours rapporté, la délimitation des frontières et le traitement instrumental. Il faut préciser que l'analyse de la production est effectuée à partir de deux types de regroupement des énoncés : un regroupement traditionnel, basé sur la classification syntaxique en discours direct et en discours indirect, puis un regroupement basé sur la classification discursive des énoncés, c'est-à-dire selon la valeur de reproduction, de *(+) narrative* à *(-) narrative*. Pour chaque type de classification des hypothèses sont posées. Les résultats pour les paramètres de fréquence fondamentale, de pause et de débit sont successivement analysés. Les énoncés regroupés selon la classification syntaxique et la classification discursive sont finalement examinés en tenant compte des facteurs sociaux.

3.1 LE CORPUS

Le corpus est tiré du corpus de discours rapportés dépouillé par VINCENT et DUBOIS (1996). L'étude de VINCENT et DUBOIS a elle-même été réalisée à partir de deux corpus d'entrevues

sociolinguistiques, le corpus Sankoff-Cedergren et le corpus de Montréal 1984 (cf. SANKOFF *et al.*:1976; THIBAULT et VINCENT:1990). Le corpus de Montréal 1984, d'où sont tirés les énoncés du corpus à l'étude, est constitué de 60 nouvelles entrevues avec des locuteurs qui avaient déjà été interviewés en 1971 pour la formation du corpus Sankoff-Cedergren sur le français parlé à Montréal. Évidemment, avec 13 ans de décalage, la classe des jeunes n'était plus représentée; 12 entrevues avec des 15 à 29 ans ont donc été ajoutées. L'ensemble de l'échantillon du corpus de 1984 ne reproduit pas la symétrie de la représentation par sexe, âge et classe sociale qui caractérisait le corpus de 1971, mais on y observe une diversité comparable (THIBAULT et VINCENT 1990:27).

Il s'agit d'un contexte conversationnel simple : celui de la relation interviewer/ interviewé inhérent à une enquête sociolinguistique. Les entrevues se sont déroulées au domicile du sujet interviewé. Elles duraient en moyenne 100 minutes pour les informateurs re-interviewés et 80 minutes pour les jeunes qui se sont ajoutés au corpus principal. Les informateurs portaient un micro-cravate afin de maximiser la qualité de l'enregistrement. Malgré les précautions prises pour assurer une qualité d'enregistrement acceptable, plusieurs éléments font en sorte que le son n'est pas idéal pour les études acoustiques (bruits de l'extérieur, bruits de musique, de lessiveuse, de lave-vaisselle; va-et-vient de la maisonnée).

3.1.1 LA QUALITÉ DE L'ENREGISTREMENT

L'enregistrement de l'entrevue doit être de qualité suffisante pour permettre une analyse acoustique fiable. Comme les entrevues cherchaient à recréer une situation de conversation naturelle, la qualité de l'enregistrement, tel que mentionné précédemment, laisse parfois à désirer. Certains locuteurs de même que certains énoncés des locuteurs retenus ont dû être rejetés.

3.1.2 LE PROFIL SOCIAL DES INFORMATEURS

Les informateurs sont sélectionnés à partir de trois facteurs sociaux: le sexe, l'âge et le groupe socioprofessionnel. Idéalement, chaque sous-groupe se compose d'au moins 2 représentants, ce qui équivaut à 16. Le groupe socioprofessionnel élevé représente les professions libérales ou patrons d'entreprise, les diplômés universitaires employés ainsi que les techniciens et contremaîtres alors que le groupe socioprofessionnel faible ou moyen représente ceux sans emploi stable, les cols bleus et les cols blancs. Les jeunes, ceux qui ne sont pas assez avancés dans leurs études (cégep non complété), sont classés selon le groupe socioprofessionnel des parents; ceux qui sont à l'université sont classés dans le groupe socioprofessionnel élevé. Chez les femmes, il n'a pas été possible de trouver deux représentantes du groupe socioprofessionnel faible ou moyen dans la catégorie 15-29 ans et du groupe socioprofessionnel élevé dans la catégorie 50 ans et plus. Parmi ces groupes, aucune locutrice ne remplissait toutes les conditions précédemment énumérées. Dans les faits, la composition sociolinguistique du corpus est telle qu'illustrée dans le *Tableau 3.1*.

Tableau 3.1 Composition sociolinguistique du corpus.

	JEUNES (n = 8) **(15-29 ans)**	**AÎNÉS** (n = 8) **(50 ans et plus)**
HOMMES (n = 8)	Groupe socioprofessionnel faible ou moyen (n = 2) loc. 99-130	Groupe socioprofessionnel faible ou moyen (n = 2) loc. 30-90
	Groupe socioprofessionnel élevé (n = 2) loc. 121-126	Groupe socioprofessionnel élevé (n = 2) loc. 73-76
FEMMES (n = 8)	Groupe socioprofessionnel faible ou moyen (n = 1) loc. 125	Groupe socioprofessionnel faible ou moyen (n = 3) loc. 15-59-108
	Groupe socioprofessionnel élevé (n = 3) loc. 123-127-131	Groupe socioprofessionnel élevé (n = 1) loc. 79

3.1.3 LES TYPES DE DISCOURS RAPPORTÉ

L'entrevue doit contenir une quantité suffisante de discours introduits par le verbe *dire*, de discours directs sans verbe *dire* et de discours indirects[1]. Cependant, il faut considérer que l'emploi du discours rapporté relève de certains facteurs individuels de même que de facteurs discursifs : l'emploi du discours direct avec un verbe introducteur est plus fréquent que sans verbe introducteur et le discours direct est beaucoup plus employé que le discours indirect. La totalité de la production des 16 locuteurs se distribue ainsi : discours direct avec le verbe *dire* (576), sans verbe *dire* (169) et discours indirect (104)[2]. Parmi ces 849 énoncés, une moyenne de 5 énoncés par type de discours a été jugée représentative de la production d'un locuteur. Toutefois, il n'a pas toujours été possible de maintenir, chez tous les informateurs, cette moyenne de 5 énoncés dans chaque type de

---

[1] Cette subdivision ne correspond pas à celle effectuée par VINCENT et DUBOIS (1996). En effet, ces dernières distribuent les discours rapportés en discours introduits par le verbe *dire* (70%), introduits par des verbes spécifiques (verbes locutoires modalisés, verbes de parole autres que *dire* et verbes d'action) (13%), ou non introduits par un verbe (17%). Étant donné que l'objectif de la présente étude consiste à examiner les relations entre la prosodie et la syntaxe, il fallait départager le discours direct et le discours indirect. Les énoncés représentant ces deux types de discours sont principalement introduits par le verbe *dire*. Le grand nombre d'occurrences des discours introduits par le verbe *dire*, la neutralité discursive de ce dernier (il est cependant parfois modalisé) et son utilisation fréquente sous une forme monosyllabique assurent un minimum d'homogénéité dans les énoncés retenus. D'un autre côté, il fallait aussi des énoncés sans verbe *dire* (DD$_{\emptyset}$). Quelques-uns de ces énoncés DD$_{\emptyset}$ (13/69) contiennent des verbes de communication plus ou moins éloignés du verbe *dire* (souvent porteurs d'attitude et la plupart du temps utilisés sous une forme plurisyllabique), mais la majorité des énoncés DD$_{\emptyset}$ ne sont introduits par aucun verbe (56/69).

[2] Cette distribution (88% de discours directs et 12% de discours indirects) correspond à celle observée pour l'ensemble du corpus VINCENT et DUBOIS (1996). En effet, sur 4860 énoncés rapportés, 4176 sont en discours direct (86%) et 684, en discours indirect (14%).

discours rapporté. L'entier du corpus de la présente étude est constitué de 215 énoncés : 94 énoncés en discours direct avec verbe *dire*, 69 en discours direct sans verbe *dire* et 52 en discours indirect. Les énoncés devaient garder un sens hors contexte tout en étant d'une longueur raisonnable (entre 4 et 30 syllabes, pour une médiane de 13 syllabes). Le *Tableau 3.2* donne la répartition des types de discours rapporté, pour chacun des locuteurs.

Tableau 3.2 Répartition des discours rapportés dans chaque type de discours et pour chaque locuteur

	D. DIRECT	D. DIRECTø	D. INDIRECT	*TOTAL*
Locuteur 18	5	5	5	15
Locuteur 30	3	2	2	7
Locuteur 89	6	4	5	15
Locuteur 73	6	3	3	12
Locuteur 76	5	3	1	9
Locuteur 79	5	6	5	16
Locuteur 90	5	0	2	7
Locuteur 99	6	5	5	16
Locuteur 108	8	6	5	19
Locuteur 121	5	5	0	10
Locuteur 123	7	4	5	16
Locuteur 125	6	6	5	17
Locuteur 126	5	5	4	14
Locuteur 127	6	5	5	16
Locuteur 130	7	5	0	12
Locuteur 131	9	5	0	14
TOTAL	94	69	52	215

Le dépouillement du corpus de discours rapportés effectué par VINCENT et DUBOIS (1996) a presque été utilisé tel quel. Le critère de base de la sélection faite par les deux auteures est l'introduction, réelle ou possible, d'un verbe de parole. Dans le présent corpus, le type discours direct sans verbe *dire* comprend une majorité d'énoncés qui ne

sont introduits par aucun verbe, leur sélection repose sur la possibilité d'introduire un verbe de parole entre le discours citant et le discours cité. Les autres énoncés, ceux introduits par un verbe autre que *dire*, contiennent des verbes, plus ou moins éloignés du verbe *dire*[3], se rapprochant plus ou moins de certaines valeurs d'attitudes évoquées dans GREVISSE.

(a) Alors Trochut levait au ciel ses mains dodues, et **larmoyait** : «Il était un pauvre homme, il avait eu si peur [...]» (Genevoix, *Raboliot*, II, 2 cité par GREVISSE (1986, no 406, Remarque c).

L'impact de ces verbes spécifiques sur la prosodie sera vérifié en départageant les résultats obtenus pour la configuration tonale à la transition des énoncés sans aucun verbe introducteur par rapport à ceux introduits par ces verbes spécifiques, autres que *dire* (cf. 3.2.1, note 7).

Les énoncés doivent surtout représenter des énoncés déclaratifs. C'est le cas de la majorité (198 sur 215). Dix-sept énoncés du discours direct présentent cependant un discours cité interrogatif, ils sont produits par 12 des 16 locuteurs. Ces discours cités interrogatifs ont été retenus afin d'observer si le type d'énoncé avait une influence sur la manière prosodique d'introduire le discours cité (cf. 3.2.1, page 60).

3.1.4 LA DÉLIMITATION DES FRONTIÈRES[4]

D'un point de vue théorique, la délimitation des frontières de l'énoncé en oral spontané demeure une question délicate (cf. OUELLON, PARADIS et

---

[3] Ces verbes sont les suivants : *agacer, appeler (téléphoner), blâmer, conseiller, décider, donner des ordres, donner sa parole, encourager, envoyer des remarques, imposer, lancer des affaires, parler, poser une question*.

[4] Pour les critères détaillés de délimitation des énoncés et du discours cité en discours direct, cf. DEMERS 1996:91-98.

DUCHESNE 1991). Dans la pratique, on ne peut que mettre en oeuvre tout l'arsenal des critères syntaxico-sémantiques et prosodiques.

En ce qui concerne la délimitation entre les paroles citées et les paroles citantes, elle est, le plus souvent, facile à faire : présence du verbe de parole dans la partie citante, changement de pronom, changement de temps verbal, changement de déictique dans la partie citée, etc. Il est cependant des cas plus ambigus liés à la présence de pauses d'hésitation et de divers marqueurs lexicaux porteurs de valeurs discursives variables. Ainsi, par exemple, comment déterminer si une hésitation, un *euh* ou un *pi* situés à la frontière discours citant/discours cité appartiennent plutôt au rapporteur (discours citant) ou au locuteur rapporté (discours cité) ? Des paramètres prosodiques, particulièrement la pause silencieuse, se sont souvent avérés efficaces pour délimiter de façon systématique les frontières litigieuses. Dans le cas de marqueurs lexicaux à valeurs discursives ambivalentes, la délimitation a aussi été parfois effectuée à partir du mouvement intonatif.

3.1.5 LE TRAITEMENT INSTRUMENTAL

Le passage du signal analogique (signal continu, ici la parole enregistrée) à un signal numérique (signal reconstitué avec un nombre fini de valeurs) requiert certaines précautions afin que le signal ne subisse pas de déformation ou de perte de valeur : une première concerne la quantification et une seconde, la fréquence d'échantillonnage. En ce qui concerne la quantification, elle est assurée par la capacité du convertisseur (ici, il est d'une capacité de 16 bits, ce qui signifie 2^{16} possibilités de mesure, soit 65 536 possibilités).

D'autre part, selon le théorème de Nyquist, la fréquence d'échantillonnage minimale, c'est-à-dire le nombre d'échantillons par seconde, doit être au moins deux fois plus grande que la plus haute fréquence du signal à échantillonner. Dans le cas présent, compte tenu de la résolution fréquencielle du Sony TC-142 et du microphone à épinglette ECM-16, les plus hautes composantes spectrales ne

dépassent pas 10 kHz. La fréquence d'échantillonnage a donc été fixée à 20 kHz.

Après avoir ainsi numérisé le signal, la composition du signal a été réduite par un filtre passe-bas ne laissant passer que les fréquences inférieures à 5 kHz. Cette opération évite l'effet d'*aliasing* (repliement des fréquences hautes vers les fréquences basses). Le signal a ensuite été décimé à la moitié de la fréquence d'échantillonnage, soit 10 kHz. Le décimage a pour principal avantage de réduire l'espace-mémoire utilisé tout en conservant suffisamment d'informations.

L'instrument de traitement ici utilisé, *Computerized Speech Lab* (CSL) de Kay Elemetrics Corp., permet d'extraire du signal les informations pertinentes. Il nécessite cependant un réglage minutieux des paramètres d'analyse en fonction des caractéristiques de chaque locuteur. L'analyseur automatique, cherchant dans un gabarit bien circonscrit de valeurs, peut ainsi donner des valeurs plus justes. Par exemple, l'échelle d'analyse de F_0 est ajustée à la fréquence usuelle de chaque locuteur, la longueur de fenêtre doit être au moins égale à la durée totale de deux périodes. On peut également ajuster le pas d'avance de la fenêtre. Un pas plus petit que la durée ou la longueur de la fenêtre a un effet de lissage dont l'importance est inversement proportionnelle à la quantité de nouvelles données introduites.

Il arrive que l'extraction automatique fournissent des valeurs insolites (soit des valeurs nulles soit des valeurs représentant le double ou la moitié de la valeur réelle). Toutes les valeurs de F_0 ont été révisées visuellement à l'aide du spectrogramme à bande étroite puis ajustées au besoin à partir des valeurs environnantes ou du calcul manuel de la fréquence (avec la formule conventionnelle $F = 1/T$). Les valeurs révisées sont finalement soumises à un traitement de lissage mis au point par Martin Brousseau et modifié par Claude Paradis.

Les valeurs lissées accompagnées du déroulement dans le temps ont été transférées dans le programme Excel. La valeur de durée

correspondant à la durée syllabique maximale dans l'énoncé (le plus souvent autour de 400 ms) est inscrite à côté de la valeur de F_0 du début de l'énoncé (certes les énoncés débutant par une consonne sourde ne permettent pas une délimitation précise) puis à côté de la valeur de F_0 correspondant à la fin de chacune des syllabes. Il est ainsi possible d'obtenir un graphique représentatif de la durée de chacune des syllabes en même temps que de la courbe mélodique de l'énoncé, comme dans la *Figure 3.1*, présentée ci-après (cf. section 3.1.2).

3.1.6 EN SOMME

La présente étude s'appuie donc sur un corpus oral spontané de 215 énoncés représentatifs des types de discours rapporté, discours direct introduit par le verbe *dire* (DD), discours direct sans verbe *dire* (DDø) et discours indirect (DI). Les énoncés proviennent de 16 locuteurs choisis en vue de représenter les extrêmes de trois facteurs sociaux, le sexe, l'âge et le niveau socioprofessionnel. Les principes généraux de délimitation des énoncés suivent les règles syntaxiques, sémantiques et prosodiques habituelles. La délimitation fine des discours cités est le plus souvent fondée sur des critères prosodiques, particulièrement sur la pause silencieuse. Le traitement instrumental est effectué à l'aide du logiciel CSL.

3.2 PROSODIE ET SYNTAXE

L'analyse de la production est d'abord faite en fonction de la perspective syntaxique : existe-t-il une différence prosodique entre les discours direct et indirect produits spontanément ?

Les hypothèses

À partir des études consultées, les hypothèses suivantes sont posées :

1. Compte tenu que les grammaires (GREVISSE: 1986 et ARRIVÉ *et al.*: 1986) associent la prosodie du discours indirect à celle de l'énoncé déclaratif, que des études discursives (AUTHIER: 1978; MAINGUENEAU: 1986) relient plus la manifestation des attitudes au lexique (le verbe de parole utilisé) qu'à la prosodie en discours indirect, que l'étude prosodique de LÉANDRI (1993) sur le discours lu montre une intonation descendante en discours indirect (intonation courante de l'énoncé déclaratif), **l'hypothèse suivante est posée : la prosodie du discours indirect s'apparenterait à celle de l'énoncé déclaratif.**

2. Compte tenu que les études consultées reconnaissent, bien qu'à des degrés divers, la place importante de la prosodie dans la caractérisation du discours rapporté, particulièrement en discours direct,
l'hypothèse suivante est posée : la prosodie du discours direct se distinguerait de celle de l'énoncé déclaratif.

3. Compte tenu que des études sur la distinction entre la lecture oralisée et la parole spontanée (dont MEHTA et CUTLER: 1986; BLAUUW: 1995) observent que la prosodie de l'oral est moins stéréotypée que la prosodie du lu, observation précisée par de GAULMYN (1992) à l'effet que les

marques lexicales, morphologiques et gestuelles suppléent régulièrement à la prosodie du discours rapporté oral spontané,
l'hypothèse suivante est posée : la prosodie des discours rapportés en oral spontané aurait une tendance à la fluctuation, c'est-à-dire que les deux premières hypothèses ne se vérifieraient que dans une certaine mesure.

4 Compte tenu que dans le corpus à l'étude, comme dans celui de de GAULMYN (1992), il arrive souvent que des discours cités non introduits par un verbe de parole débutent par un élément lexical (comme *ah, bon, ben*, etc.),
l'hypothèse suivante est posée : l'absence des uns serait comblée par la présence des autres et en conséquence la prosodie n'en serait pas considérablement affectée.

Les résultats

Les éléments prosodiques pressentis et/ou étudiés dans les études antérieures comme marqueurs du discours direct sont l'intonation, la pause et le débit; l'intensité n'est pas évoquée. Le paramètre de l'intensité n'est pas non plus traité dans la présente étude, non pas qu'il soit jugé sans effet possible sur le marquage du discours rapporté en parole spontanée, au contraire[5]. Malheureusement, la qualité acoustique du corpus n'a pas permis une étude fiable de ce paramètre. En effet, trop de signaux étaient saturés.

---

5 Voir à ce propos GUAITELLA (1991:212) : «Nous pensons que l'intensité joue un rôle de premier plan pour l'intonation en parole spontanée [...] Nous pensons comme I. FÓNAGY (1987) que les variations d'intensité augmentent les capacités polysémiques de l'intonation. BOURDON (1892) a montré que l'intonation pouvait être une métaphore de l'intensité conditionnée socialement; ainsi, suivant le milieu social, on pourrait se permettre «d'élever la voix» (intensité), ou seulement «d'élever le ton» (mélodie).»

Afin de mettre en évidence les caractéristiques prosodiques de chaque type de discours, les résultats sont présentés parallèlement pour le discours direct avec le verbe introducteur *dire* (DD), puis pour le discours direct sans verbe *dire* (DD∅) et finalement, pour le discours indirect (DI).

3.2.1 LA FRÉQUENCE FONDAMENTALE

Le paramètre prosodique majeur qu'est la fréquence fondamentale n'est pas d'un traitement facile : l'oreille a du mal à en saisir toutes les nuances, la machine, bien que de plus en plus précise, s'égare parfois si les changements sont trop brusques, si l'articulation est plus ou moins nette, si l'enregistrement est plus ou moins clair. Les changements n'ont pas toujours de signification linguistique. Ils peuvent être attribuables à la microprosodie ou à des ratés de l'appareil phonatoire — certains cas de *creaky voice* par exemple. Les seuils de perception ne font pas véritablement consensus ('t HART, COLLIER et COHEN 1990 29-33), le point de prise de mesure non plus (voyelles ou points-cibles), etc. Malgré cette zone générale de flottement, si l'on peut dire, lorsque la même méthodologie, clairement établie, est appliquée à un corpus suffisamment important, les tendances dégagées devraient pouvoir témoigner de l'ensemble de la réalité prosodique.

Dans le cas présent, le problème du «quoi mesurer» était de taille : les énoncés ont une structure syntaxique différente, une longueur différente, ils sont prononcés par des locuteurs différents, dans des contextes différents, ils n'ont qu'une caractéristique commune : ils représentent tous du discours rapporté. Faut-il, de ce fait, ne comparer que ce qui se passe à la transition entre le discours citant et le discours cité et dans ce cas, ne s'intéresser qu'aux deux syllabes à la transition, ou plutôt élargir aux deux groupes intonatifs à la transition ? Est-il valable, malgré les différences évoquées ci-haut, de comparer également l'allure générale du discours citant, celle du discours cité, puis celle de l'énoncé dans son ensemble ? Chacun des domaines suivants a été examiné : le discours citant, le *groupe intonatif avant la transition*, la transition entre le

discours citant et le discours cité, le *groupe intonatif après la transition*, le discours cité et le profil global de l'énoncé.

Pour chaque type de discours, chaque domaine retenu est considéré selon sa pente : montante, descendante ou plate. Comme le domaine de la transition relève d'un écart intersyllabique, le seuil de perception de l'intervalle mineur (2DT) proposé par MERTENS (1993:23) (cf. section 2.2.1) est retenu.

En revanche, les domaines du *groupe intonatif avant la transition* et du profil requéraient un seuil pour le mouvement de F_0. À ma connaissance, il n'existe pas véritablement de seuil pour la perception d'un mouvement de F_0 plus étendu que celui entre deux syllabes, si ce n'est celui proposé par 't HART (1981:820), qui s'applique cependant à la différence d'ampleur entre deux mouvements intonatifs orientés dans le même sens. Dans ce cas, le seuil de la *just noticeable difference* (jnd) pour un mouvement de F_0 est de 2 DT, mais dans la conversation courante, ce seuil est haussé à 3 DT (p. 821). Ces seuils ne s'appliquent pas au type de mouvement intonatif ici observé. Ceci dit, un seuil de 2 DT, puis de 3 DT, sera arbitrairement fixé à des fins purement descriptives. Dans un premier temps, toutes les données ont été traitées avec le seuil de 2 DT, dans un deuxième temps, les tendances du *groupe intonatif avant la transition* et du profil global de l'énoncé ont également été traitées avec le seuil de 3 DT afin de vérifier si les tendances en étaient affectées. Les graphiques qui suivent montrent les résultats obtenus avec l'application du seuil de 2 DT, mais les résultats statistiques sont données pour le seuil de 2 DT et pour le seuil de 3 DT.

Ainsi un écart supérieur ou égal à +2 DT (puis à +3 DT) correspond-il à une configuration tonale montante, un écart supérieur à -2 DT (puis à -3 DT), à une configuration tonale descendante, un écart entre ± 2DT (puis ± 3 DT), à une configuration tonale plate. Les résultats sont présentés en fréquence relative (nombre de cas sur 100) et en écarts médians (DT). Le khi deux (χ^2) est appliqué afin de vérifier l'interdépendance entre la fréquence d'une configuration tonale et le type

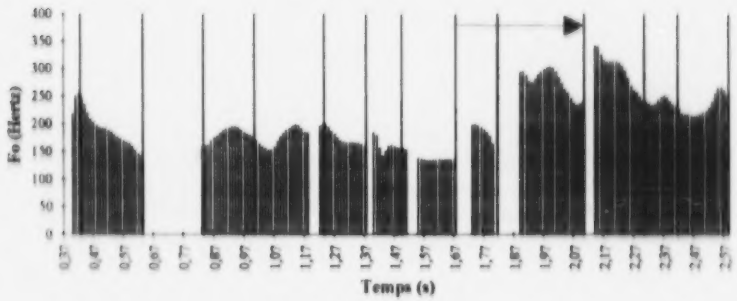
de discours de même qu'entre la fréquence de la cooccurrence des domaines (patrons intonatifs) et le type de discours rapporté[6]. Le degré de signification des variables continues est mesuré à l'aide du test Mann-Whitney non pairé. La valeur de p est donnée selon les valeurs standard reconnues :

$$p < 0{,}001,\ p < 0{,}01,\ p < 0{,}05.$$

Toutefois, les valeurs de p entre 0,05 et 0,1 sont mentionnées comme indicatrices d'une certaine tendance. Il va sans dire que celle-ci demande à être traitée avec grande prudence cependant.

À partir de la *Figure 3.1*, on peut illustrer la façon dont les mesures de F_0 ont été prises pour chacun des domaines retenus.

[6] Bien sûr le test de khi deux est appliqué à partir des fréquences absolues. Toutefois, comme le nombre d'énoncés varie d'un type de discours à l'autre, l'illustration est présentée en fréquence relative afin de faciliter les comparaisons.



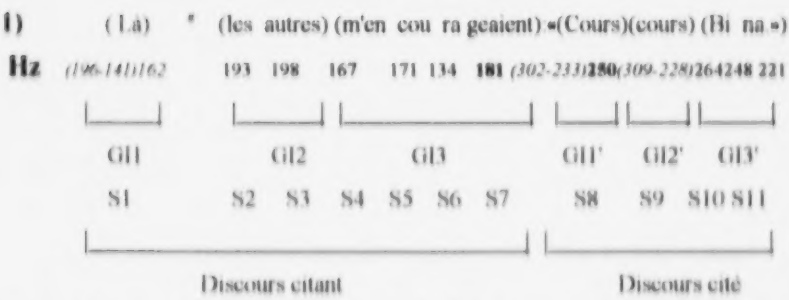


Figure 3.1 Courbe intonative de *Là " les autres m'encourageaient : «Cours cours Bina »* (5609ɔ-108).

Il est peut-être utile de rappeler d'abord que la lecture des valeurs de F_0 se fait aux deux tiers de la durée vocalique. Dans cet exemple, le *groupe intonatif avant la transition* (GI3), qui s'étend de 167 à 181 Hz, est égal à un écart de 1,77 DT; ce *groupe intonatif avant la transition* est considéré comme plat. La mesure de la transition (indiquée par la flèche) se fait entre la dernière syllabe du discours citant et la première syllabe du discours cité, soit entre la syllabe 7 (S7), 181 Hz, et la syllabe 8 (S8), 250 Hz, ce qui correspond à un écart de 5,54 DT; la transition est donc considérée comme montante. Le profil global de l'énoncé est mesuré par l'écart entre la première syllabe du discours citant (S1), 162 Hz, et la dernière syllabe du discours cité (S11), 221 Hz, ce qui correspond à un écart de 5,38 DT; le profil est donc considéré comme montant.

Le groupe intonatif avant la transition

Tel que défini précédemment, le discours citant représente la partie du discours qui précède les paroles citées. Il peut comprendre un ou plusieurs groupes intonatifs. Celui qui précède la transition est porteur de caractéristiques fréquencielles distinctives des types de discours rapporté.

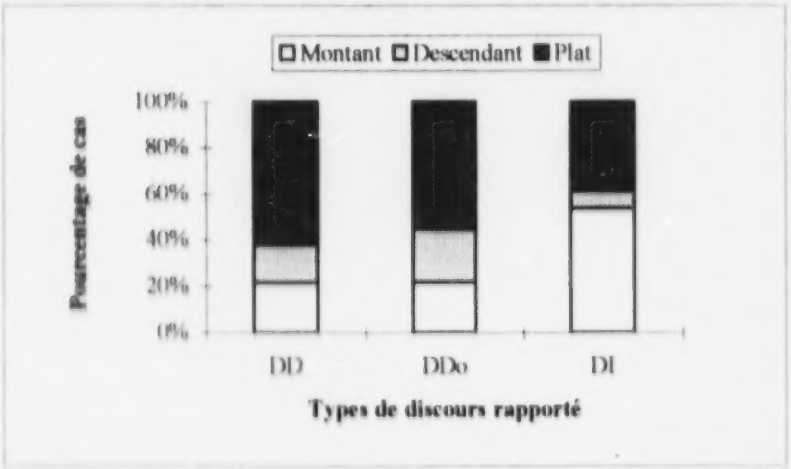


Figure 3.2 *Groupe intonatif avant la transition.* La fréquence relative des cas montants, descendants et plats, pour le discours direct avec le verbe *dire* (DD), le discours direct sans verbe *dire* (DDo) et le discours indirect (DI).

Les résultats présentés dans la *Figure 3.2* indiquent que le comportement du *groupe intonatif avant la transition* présente la même tendance dans les deux sous-catégories de discours direct, il est la plupart du temps plat (62% et 55%) tandis qu'en discours indirect, le *groupe intonatif avant la transition* est le plus souvent montant (54% des cas). ($p < 0,001$, que le seuil soit de 2 DT ou de 3 DT).

La transition discours citant/discours cité

Le passage du discours citant au discours cité est ici observé en considérant l'écart entre la valeur fréquencielle de la dernière syllabe du discours citant et celle de la première syllabe du discours cité.

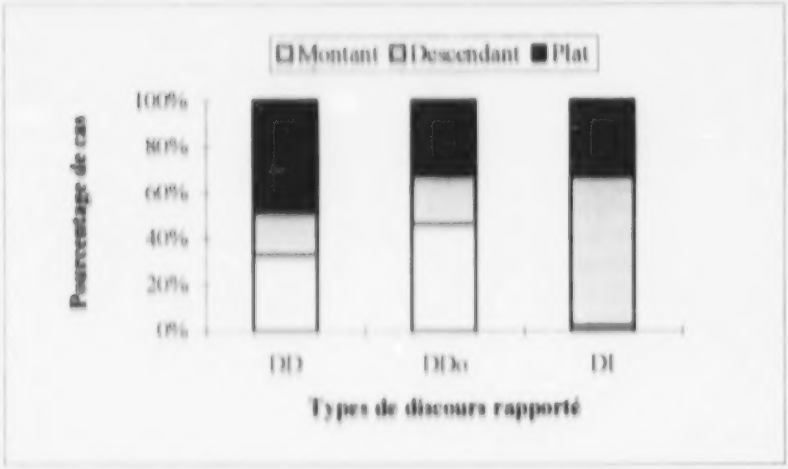


Figure 3.3 Transition. La fréquence relative des cas montants, descendants et plats, pour le discours direct avec le verbe *dire* (DD), le discours direct sans verbe *dire* (DDs) et le discours indirect (DI).

L'examen de la transition discours citant/discours cité, *Figure 3.3*, montre, d'une part, une opposition nette entre le discours indirect et les deux sous-catégories de discours directs : en discours indirect, la transition est la plupart du temps descendante (65% des cas), ce qui est beaucoup moins fréquent en discours direct (19% et 22%). De plus, le discours indirect présente très rarement une transition montante (2%); ($p < 0,001$).

L'importance du domaine de la transition justifie qu'on y regarde de plus près. Les deux sous-catégories de discours direct présentent assez

régulièrement des transitions montantes (respectivement, 33% et 46%)[7]. On remarque d'autre part que les trois types de discours rapporté présentent assez souvent des transitions plates : 48% en discours direct avec verbe *dire* (DD), 32% en discours direct sans verbe *dire* (DD∅) et 33% en discours indirect (DI). Dans des proportions différentes, la transition descendante peut donc être commune au discours direct et au discours indirect. Il s'agit donc là de «zones d'intersection» qui demanderont à être examinées en cooccurrence avec les autres domaines et les autres paramètres prosodiques afin de voir s'il existe, à un deuxième niveau, des caractéristiques distinctives. En bref, seule la transition montante semble à peu près exclusive au discours direct. La transition descendante, bien que fortement liée au discours indirect, ne lui est pas exclusive alors que la transition plate est commune aux trois types de discours rapporté.

Le profil global de l'énoncé

La configuration tonale début/fin du discours cité n'est pas distinctive des types de discours rapporté; celle début/fin de l'énoncé l'est-elle davantage ? On pourrait s'attendre à ce que ce soit le cas. En effet, la neutralité discursive prêtée au discours indirect devrait se rapprocher d'une configuration tonale générale montante/ descendante, d'autant plus que le *groupe intonatif avant la transition* est de tendance montante et que la transition est le plus souvent descendante. D'autre part, les différentes valeurs discursives du discours direct risqueraient de montrer un profil général plus modulé, avec une tendance vers une configuration tonale générale d'énoncé montante.

---

[7] On se souvient que la classification des discours directs sans verbe *dire* contenait des verbes s'approchant plus ou moins de ce qu'est un verbe de parole (cf. section 3.1.3). La différence de fréquence relative (%) entre DD et DD∅ du point de vue de la configuration tonale de la transition a été comparée selon que la catégorie DD∅ contient ou ne contient pas de verbes de communication; l'analyse statistique s'est avérée non significative.

Toutefois, il faut bien admettre que la mesure est fragile compte tenu, entre autres, de la longueur très variable des énoncés. De plus, ce domaine correspond à la ligne de déclinaison dont la mesure précise est loin de se limiter à la différence entre le la hauteur de la première syllabe et la hauteur de la dernière syllabe de l'énoncé (cf. à ce propos l'étude de LIEBERMAN, KATZ, JONGMAN, ZIMMERMAN et MILLER (1985). Les résultats obtenus pour ce domaine ne présentent donc qu'une valeur toute relative.

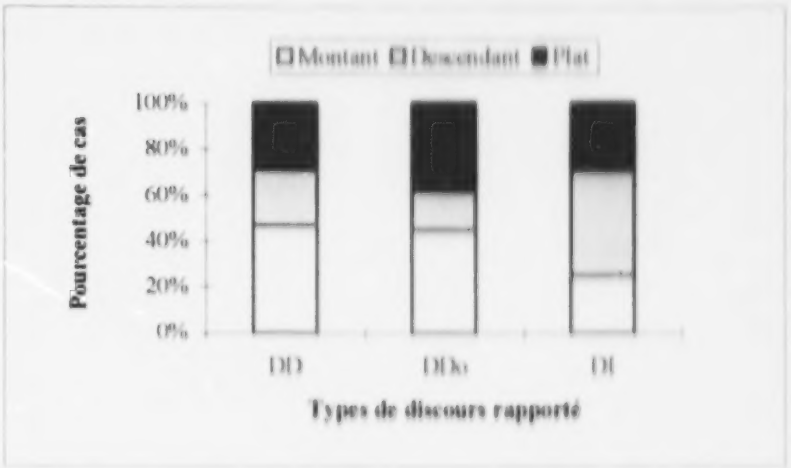


Figure 3.4 Profil global de l'énoncé. La fréquence relative des cas montants, descendants et plats, pour le discours direct avec le verbe *dire* (DD), le discours direct sans verbe *dire* (DDø) et le discours indirect (DI).

Il ressort de la *Figure 3.4* que le profil global de l'énoncé en discours direct est effectivement plus souvent montant (47% et 45%) qu'en discours indirect (25%) et inversement, plus fréquemment descendant en discours indirect (46%) qu'en discours direct (respectivement 24% et 17%). Le profil global plat est cependant commun à tous les types de discours rapporté, dans des proportions

non négligeables. (29%, 38%, 29%). Toutefois, les différences sont significatives; ($p < 0,01$, que le seuil soit de 2 DT ou de 3 DT).

Pour l'ensemble des domaines, il apparaît que les tendances se maintiennent aussi bien lorsque le seuil de perception est fixé à 3 DT que lorsqu'il est fixé à 2 DT. Désormais, les données seront traitées à partir du seuil de 2 DT.

On se souvient que le corpus des discours directs contient 17 énoncés interrogatifs. L'analyse statistique montre que, pour le *groupe intonatif avant la transition*, la transition et le profil global de l'énoncé, la différence de fréquence relative entre DD et DDø selon la présence ou l'absence des énoncés interrogatifs est non significative.

L'ampleur des écarts

On vient de voir qu'en ce qui concerne la fréquence relative, ce sont les domaines du *groupe intonatif avant la transition*, de la transition et du profil global de l'énoncé qui ressortent comme représentatifs des tendances contrastives entre discours direct (désormais /discours direct/, ce qui désigne les deux sous-catégories de discours direct réunies[8]) et discours indirect. Qu'en est-il si on considère l'ampleur des écarts dans chacun des domaines ?

Les résultats précédents sont confirmés :

a) Le *groupe intonatif avant la transition* discours citant/discours cité présente une médiane montante en discours indirect (2,62 DT) alors que cette montée est nulle en /discours direct/ (0 DT), ($p < 0,001$)

[8] Aucun domaine ne présente des caractéristiques fréquentielles qui permettent véritablement de distinguer le discours direct syntaxiquement «fort» (DD) du discours direct syntaxiquement «faible» (DDø).

b) La transition entre discours citant/discours cité a la tendance inverse, c'est-à-dire qu'elle est descendante en discours indirect (-4 DT), ce qui n'est pas le cas en /discours direct/ (0,78 DT), ($p < 0,001$).

c) Le profil global de l'énoncé est positif en /discours direct/ (1,54 DT) et négatif en discours indirect (-0,7 DT); ($p < 0,01$). Le fait que la médiane des configurations tonales de ce domaine se situe en-dessous des seuils de perception utilisés, de même que les réserves émises précédemment, amène à penser que la distinction entre les deux types de discours rapporté par le profil est peu probablement perceptible.

La cooccurrence des configurations tonales

Il apparaît maintenant opportun d'observer dans quelle mesure la cooccurrence des configurations tonales *groupe intonatif avant la transition*/ transition/ profil global de l'énoncé peut préciser les caractéristiques mentionnées ci-dessus.

Seules les combinaisons dont le nombre d'énoncés est ≥ 5% d'une catégorie de discours sont retenues. Le *Tableau 3.3* illustre les patrons tendanciels pour chacun des grands types de discours rapporté.

Tableau 3.3 Patrons tendanciels (montant, descendant ou plat) dans les domaines significatifs : *groupe intonatif avant la transition*, transition et profil global de l'énoncé, en /discours direct/ (/DD/) et en discours indirect (DI).

	GI av. transition	Transition	Profil de l'énoncé
/DD/	Plat (Desc.)	Montant	Montant*
DI	Montant	Desc.	(Desc., Plat, Montant*)
	Plat	Desc.	Desc.

*Les résultats statistiques sont légèrement supérieurs au seuil de 0,05.

Le discours indirect est donc significativement lié à un *groupe intonatif avant la transition* montant en cooccurrence avec une

transition descendante et un profil global descendant (MD-D), plat (MD-P) ou montant (MD-M*). Le *groupe intonatif avant la transition* plat et la transition descendante s'associent plutôt à un profil descendant (PD-D). En revanche, le /discours direct/ est significativement caractérisé par un *groupe intonatif avant la transition* plat ou descendant, cooccurrent à une transition montante et à un profil montant (PM-M* et DM-M*). Il est à remarquer que la cooccurrence avec le profil montant, aussi bien en discours direct qu'en discours indirect, est un indice plus fragile.

Les énoncés à transition plate relativement nombreux, tant en /discours direct/ (40% des énoncés) qu'en discours indirect (33%), n'ont pas été plus distinctifs lorsqu'examinés en fonction du *groupe intonatif avant la transition* et du profil de l'énoncé. La zone d'intersection à la transition est donc trop importante pour que l'on puisse prédire le type de discours rapporté par les seules caractéristiques fréquencielles.

De plus, il faut rappeler que les patrons complets (*groupe intonatif avant la transition* /transition/profil global de l'énoncé) ne correspondent qu'à 16% de l'ensemble des /discours directs/ par rapport à 60% des discours indirects. Il s'agit là d'une démonstration claire que les qualités fréquencielles du discours indirect sont beaucoup plus stables que celles du /discours direct/, mais aussi que de façon générale, la correspondance entre la syntaxe et l'intonation est loin d'être catégorique. D'autres paramètres prosodiques, seront examinés : la pause et le débit, mais il semble d'abord nécessaire d'examiner l'effet sur la fréquence fondamentale d'éléments spécifiques au /discours direct/.

3.2.2 L'EFFET DES «MARQUEURS TRANSITIONNELS» DU DISCOURS DIRECT SUR F_0

Dès que l'on examine le moindrement la construction des /discours directs/, on se rend compte que les paroles du discours cité débutent assez fréquemment par des particules de différents genres,

regroupées sous l'appellatif «marqueurs d'attaque»[9]. *Ah, ah ben, ah là, oh, aye, ben, ben écoute, ben là, ben tsé, bon, coudon, oui, non, va donc, mon-dieu* et *pi* représentent la majorité des marqueurs d'attaque, le reste étant principalement des vocatifs, *maman, François, Gigi*, etc., des sacres, *criss*, des jurons, *crime*, ou une salutation, *bonjour*. Ainsi retrouve-t-on :

(2) (Comme le vin) (: «**Ah**) (celui-là) (il est bon) (celui-là il est bon.») (6143ø-131)

(3) (Ben j'étais là) (: «**Crime**) (c'est facile.») (Tsé.) (6048ø-127)

(4) (Elle dit) (: « **Gigi**) # (on y va.») (4657-59)

Il est également assez fréquent de trouver des pauses entre le discours citant et le discours cité. Ces pauses peuvent aussi bien être des pauses silencieuses[10] que des pauses d'hésitation[11]. Ainsi les exemples (5) et (6) :

(5) (C'est une fille) (qui était pas rodée) (pi tout 'suite) # (: «Ces deux-là) (dehors.») (4065ø-15)

(6) (Elle dit &[ruh:375 ms]) # (: «Tu as l'air d'un gars) (qui a bu) (en fin de semaine.») (5877-126)

---

9 Le terme «marqueur d'attaque» représente «l'ensemble des manifestations verbales que le locuteur émet lorsqu'il prend la parole : *ah, oui, ben ,mais*, etc.» (VINCENT et DUBOIS 1996:21), ce que VICHER et SANKOFF (1989) appellent *pre-sentential turn-openings*. Certains de ces marqueurs lexicaux (*pi* et *tsé* par exemple) peuvent jouer d'autres rôles discursifs que celui inhérent à la prise de parole. Dans ce cas, ils ne sont pas des marqueurs d'attaque et ne font évidemment pas partie du discours cité.

10 Par pause silencieuse, il faut entendre l'absence d'émission vocale d'une durée ≥supérieure ou égale à 250 ms, tel que proposé par GOLDMAN-EISLER (1968); c'est le seuil utilisé depuis par la majorité des chercheurs.

11 Par pause d'hésitation, il faut entendre toute émission vocale supérieure ou égale à 200 ms (MERTENS: 1987; GUAITELLA:1992), qui ne fait pas sens. *Euh* est de loin la plus fréquente, mais on retrouve également des sons comme *m:, b:, v:, chpi:, vaj:* etc.

Les marqueurs d'attaque ainsi que les pauses à la transition sont ici désignés comme des «marqueurs transitionnels» du /discours direct/[12]. Ces «marqueurs transitionnels» peuvent agir en conjonction, c'est-à-dire que l'on peut en retrouver deux, voire trois, à la transition[13]. Dans l'exemple (7), il y a conjonction de la pause silencieuse et du marqueur d'attaque *ah*.

(7) (C'est les autres)(qui vous voient)(mettre le billet) #
(: «**Ah**)(attends donc.») (5483⌀-99)

La présence de ces marqueurs soulève deux questions :

1) Est-ce que les «marqueurs transitionnels» jouent un rôle dans la distinction des deux sous-catégories de discours direct, le discours direct avec verbe *dire* et le discours direct sans verbe *dire* ?
2) Est-ce que les valeurs fréquencielles à la transition sont différentes lorsqu'il y a un marqueur d'attaque ou une pause ?

---

12 Dans le corpus étudié, ces éléments constituent une caractéristique absolue du discours direct puisque, en aucun cas, des pauses et des *ah*, *ben*, *coudon*, etc ont été relevés avant le *que* du discours indirect. Toutefois, bien qu'aucun cas n'ait été relevé, il est plausible qu'une pause, silencieuse ou d'hésitation, puisse se glisser avant le *que*. D'ailleurs dans le corpus de MERTENS (1987:44), on en trouve un exemple :
*(Roland Barthes # ils sont) (nombreux) (ceux qui affirment # **qu'**il est difficile) # (de vous situer.)*
Il semble cependant que ce soit rare. Il en est de même pour les *ah*, *ben*, *coudon*, etc. Comme ils constituent un indice de prise de parole et que le locuteur «rapporté» est évincé du discours indirect, il est également rare qu'on en retrouve après le *que*. Cependant, quelques cas ont été relevés dans le corpus de VINCENT et DUBOIS (1996).

13 En discours direct avec verbe *dire* (DD), il y a conjonction pause d'hésitation/pause silencieuse (n=3), pause silencieuse/marqueur d'attaque (n=2) et pause d'hésitation/marqueur d'attaque (n=1); en discours direct sans verbe *dire* (DD⌀), la conjonction pause silencieuse/marqueur d'attaque est la plus courante (n=8), les associations pause d'hésitation/pause silencieuse (n=3), pause d'hésitation/marqueur d'attaque (n=2) et pause silencieuse/pause d'hésitation/ marqueur d'attaque (n=1) constituent les autres cas.

Le *Tableau 3.4* présente la distribution des «marqueurs transitionnels» dans le discours direct avec verbe *dire* (DD) et dans le discours direct sans verbe *dire* (DD⌀).

Tableau 3.4 Distribution des «marqueurs transitionnels» (MT) dans les deux sous-catégories de discours direct (/DD/) : en discours direct avec verbe *dire* (DD) et en discours direct sans verbe *dire* (DD⌀). (MA) désigne le marqueur d'attaque; (Ps), la pause silencieuse; (Ph), la pause d'hésitation; (Conjoints) les énoncés dans lesquels il y a plus d'un «marqueur transitionnel».

		MT			
DD/ avec n =163	**MT** n=91 (56%)	**MA** n=68 (42%)	**Ps** n=29 (18%)	**Ph** n=14 (9%)	**(Conjoints)**[14] n=20 (12%)
DD n=94	43 (46%)	32 (34%)	10 (11%)	7 (7%)	6 (6%)
DD⌀ n=69	48 (70%)	36 (52%)	19 (28%)	7 (10%)	14 (20%)

Il apparaît que les «marqueurs transitionnels» se manifestent dans une proportion significativement plus importante lorsque le discours direct est sans verbe *dire* (DD⌀), 48 cas sur 69 énoncés (70%) par rapport à 43 cas sur 94 énoncés (46%) lorsque le discours cité est précédé du verbe *dire* (DD), ($p < 0,01$). Il en est de même pour chacun des «marqueurs transitionnels» pris séparément (MA : $p < 0,05$; Ps : $p < 0,01$), sauf pour la pause d'hésitation (n.s.). Les cas de pauses d'hésitation étant trop peu nombreux, on ne peut ni en tirer de conclusions ni en faire un examen plus approfondi.

On s'en tiendra donc aux marqueurs d'attaque et aux pauses silencieuses. La fréquence relative et l'ampleur des écarts des

[14] Les énoncés contenant des cas conjoints (plus d'un «marqueur transitionnel» à la transition) sont bien sûr soustraits du nombre total d'énoncés contenant des «marqueurs transitionnels».

tendances, dans les trois domaines significatifs, ont été examinées pour les deux sous-catégories de discours direct. Les énoncés sans marqueur d'attaque ont été comparés aux énoncés avec marqueur d'attaque; de même les énoncés sans pause silencieuse à la transition ont été comparés aux énoncés avec pause silencieuse à la transition.

Le marqueur d'attaque. En bref, l'effet du marqueur d'attaque se manifeste à deux niveaux :

1) Au niveau syntaxique. Le marqueur d'attaque supplée assez souvent au verbe *dire* : dans 52% des discours directs sans verbe *dire*, c'est un marqueur d'attaque qui ouvre le discours cité.
2) Au niveau intonatif. Le marqueur d'attaque ne modifie pas significativement la ligne mélodique de l'énoncé. Cependant, il apparaît intéressant de retenir que la montée de la transition est progressive lorsque le discours cité est introduit par un marqueur d'attaque polysyllabique[15]. Reste à savoir si cette montée progressive est perceptible, au même titre que la transition montante, comme une caractéristique du discours direct.

La pause silencieuse. Les pauses silencieuses sont en nombre moins important (18% des /discours directs/ en contiennent) que les marqueurs d'attaque (42% des /discours directs/) à la transition discours citant/discours cité. L'effet de la pause silencieuse se manifeste également à deux niveaux :

---

15 En effet, si la plupart des marqueurs d'attaque sont monosyllabiques (49/68), il s'en trouve tout de même 28% (19/68) qui sont polysyllabiques. On remarque qu'une majorité de marqueurs polysyllabiques présentent des montées progressives jusqu'à la fin du marqueur (62% des cas), Deux cas dans le présent corpus (5612-108 et 6095-130) présentent une montée **après** le marqueur d'attaque polysyllabique; dans les deux cas, il s'agit du marqueur polysyllabique *ah ben*.

1) Au niveau syntaxique. La pause silencieuse semble également suppléer au verbe *dire* : 28% dans les discours directs sans verbe *dire* par rapport à 11% dans les discours directs avec verbe *dire*.

2) Au niveau intonatif. Il paraît y avoir une différence significative entre les énoncés sans pause silencieuse et ceux avec pause silencieuse, du moins du point de vue de la fréquence relative : lorsqu'il y a une pause, le nombre de transitions plates diminue en discours direct avec verbe *dire* (DD) et le nombre de transitions montantes augmente en discours direct sans verbe *dire* (DDø); ($p < 0,05$) . D'un autre côté, les écarts médians montrent que la pente de la transition après pause silencieuse a tendance à monter tandis que celle du profil a tendance à s'aplatir. Ces dernières tendances ne sont cependant pas validées statistiquement. **On ne peut donc ici affirmer que la pause silencieuse modifie la ligne mélodique de l'énoncé**.

Dans une étude prosodique du discours rapporté (phrases lues), LEANDRI (1993:155) affirme que la présence d'une pause silencieuse à la transition discours citant/discours cité entraîne deux stratégies différentes : dans un cas, les éléments de chaque côté de la pause sont au même niveau ("D" 4); dans l'autre cas, il y a montée avant la pause puis baisse au début du discours cité ("D" 5). La *Figure 3.5* reproduit les deux stratégies présentées par LEANDRI (1993:155).

("D"4 & "D"5

Figure 3.5 Pause silencieuse. Les deux stratégies possibles lorsqu'il y a pause silencieuse à la transition en discours direct avec verbe introducteur ("D")

Dans le présent corpus, les énoncés avec pauses silencieuses à la transition ne présentent pas les stratégies observées par LÉANDRI. On aura plutôt, à l'occasion, une montée après la pause, comme dans la *Figure 3.6*.

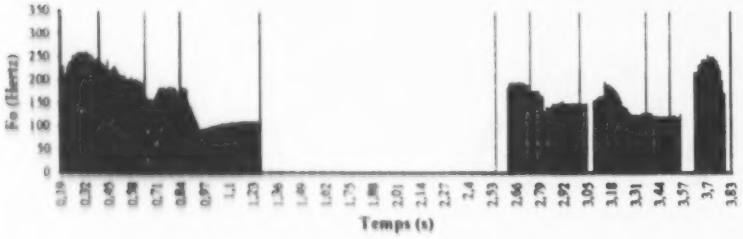


Figure 3.6 Pause silencieuse. Courbe intonative de *Là là j'ai dit # : «Faut 'je me trouve une job.»* (5309-90)

LÉANDRI se réfère à une étude précédente, non citée; on ignore le nombre d'énoncés qui sont à la base de ses observations. Quoi qu'il en soit, le corpus de LÉANDRI est un corpus lu, ce qui pourrait être à l'origine des différences observées.

Dans la perspective d'une analyse de la perception, il sera important d'examiner dans quelle mesure les effets observés agissent sur la perception du discours direct et sur sa distinction avec le discours indirect.

3.2.3 LA PAUSE

La pause a d'abord été examinée en fonction de son effet sur la ligne mélodique du discours rapporté. Cependant, elle constitue en elle-même un paramètre prosodique. Elle sera maintenant étudiée en fonction de sa fréquence d'occurrence à la transition et de sa durée dans les différents types de discours rapporté[16].

La fréquence d'occurrence des pauses

La fréquence d'occurrence des pauses a déjà été mentionnée pour chacune des sous-catégories de discours direct (*Tableau 3.4*). Elle est ici présentée (*Tableau 3.5*) pour l'ensemble des discours directs (/DD/) par rapport au discours indirect.

Tableau 3.5 Les pauses. Fréquence absolue et relative des pauses à la transition entre le discours citant et le discours cité pour l'ensemble des discours directs (/DD/) ainsi que pour les discours indirects (DI).

	/DD/ (n=163)	DI (n=52)
Pauses silencieuses	29 (18%)	0
Pauses d'hésitation	14 (9%)	0

[16] Rappelons que la pause silencieuse est définie comme l'absence d'émission vocale, d'une durée supérieure ou égale à 250 ms (GOLDMAN-EISLER:1968) alors que la pause d'hésitation est définie comme une émission vocale qui ne fait pas sens, d'une durée supérieure ou égale à 200 ms (MERTENS:1987; GUAÏTELLA:1992).

L'absence absolue de pause entre le discours citant et le discours cité en discours indirect par rapport à sa présence occasionnelle en /discours direct/ est reliée à la valeur syntaxique attribuée aux deux grandes catégories de discours rapporté : le discours indirect est construit par la subordination du discours cité et de ce fait n'entraîne pas d'interruption pausale alors que le discours direct est constitué de deux indépendantes qui peuvent, à l'occasion, être clairement séparées par une pause silencieuse et/ou d'hésitation, les pauses silencieuses étant plus fréquentes que les pauses d'hésitation.

Par ailleurs, contrairement au paramètre prosodique de la fréquence fondamentale, le paramètre prosodique de la pause, du moins de la pause silencieuse, semble souvent lié à la «force» syntaxique du discours. Effectivement, l'apparition de la pause silencieuse est deux fois plus fréquente lorsque la structure syntaxique du discours direct est «faible», c'est-à-dire lorsque le discours cité n'est pas introduit par un verbe *dire* (DD∅). Tel que démontré précédemment (cf. *Tableau 3.4*), cette différence est significative.

La durée des pauses

Parallèlement à la fréquence d'apparition des pauses, la durée de celles-ci à la transition a été comparée à la durée de celles qui apparaissent ailleurs à l'intérieur du discours rapporté[17]. Afin de pouvoir comparer équitablement les deux distributions, toutes les durées sont prises en considération, c'est-à-dire que sont incluses les pauses silencieuses inférieures à 250 ms et les pauses d'hésitation inférieures à 200 ms, durées sous les seuils de perception retenus, tel qu'indiqué à la note 17.

---

17 La numérisation de chacun des énoncés ne permet pas de relever de façon systématique **toutes** les pauses, silencieuses et d'hésitation, qui pourraient se trouver à la fin de l'énoncé rapporté; de ce fait, les pauses relevées en fin d'énoncé n'ont pas été retenues.

Tableau 3.6 Durée des pauses (ms). Durée médiane des pauses silencieuses (Ps) et des pauses d'hésitation (Ph) à la transition puis ailleurs dans l'énoncé dans les trois types de discours rapporté : en discours direct avec verbe *dire* (DD), en discours direct sans verbe *dire* (DD∅) et en discours indirect (DI). La fréquence absolue d'apparition est entre parenthèses.

	DD (n=94)	**DD∅** (n=69)	**DI** (n=52)
Ps/transition	954 ms (11)	485 ms (23)	∅
Ps/ailleurs	465 ms (15)	272 ms (25)	501 ms (13)
Ph/transition	211 ms (12)	364 ms (9)	∅
Ph/ailleurs	276 ms (9)	278 ms (11)	233 ms (13)

D'une part, le *Tableau 3.6* montre que les pauses silencieuses ont tendance à être plus longues que les pauses d'hésitation, et ce, quelle que soit leur distribution; ces résultats confirment ceux de GUAÏTELLA (1992:123). D'autre part, les pauses à la transition sont significativement plus longues que les autres ($p < 0,05$); ce qui en soi ne surprend pas. GOLDMAN-EISLER (1968) a déjà démontré que plus la pause sépare un élément élevé dans la hiérarchie syntaxique, plus elle est longue. En discours direct avec verbe *dire* (DD), la pause d'hésitation fait ici exception : elle est plus courte à la transition qu'ailleurs (211 ms vs 276 ms); il faut préciser que (i) 5 pauses d'hésitation sur 12 sont inférieures au seuil de perception et n'ont donc pas été considérées comme «marqueurs transitionnels», (ii) mais que surtout le nombre est peu important (3 cas sur les 7 qui restent). Il est probable que ce ne soit pas surtout par sa durée que la pause d'hésitation marque la distinction discours citant/discours cité, mais beaucoup plus par sa valeur de F_0 relativement basse, effet qui semble provocateur d'une transition à la hausse.

On pourrait également s'étonner de ce que la durée des pauses silencieuses à la transition en discours direct avec verbe *dire* (DD) soit plus importante que celle en discours direct sans verbe *dire* (DD∅)

puisque, dans ce dernier cas, on serait tenté de croire qu'elles tiennent lieu de verbe de parole. Il apparaît que la présence de la pause soit plus déterminante que sa durée.

La présence de la pause silencieuse à la transition est reliée à la structure syntaxique du discours rapporté.

1) Une pause silencieuse à la transition permet de distinguer le discours direct et le discours indirect : présence occasionnelle dans le premier cas et absence absolue (du moins dans ce corpus) dans le second.
2) Une pause silencieuse à la transition permet également de distinguer le discours direct avec verbe *dire* du discours direct sans verbe *dire*: elle est significativement plus fréquente en l'absence d'un verbe *dire*.

La durée de la pause silencieuse à la transition est aussi fonction de la hiérarchie syntaxique de la phrase.

3) La durée de la pause silencieuse est significativement plus importante entre le discours citant et le discours cité qu'ailleurs à l'intérieur de l'énoncé de discours rapporté.

3.2.4 LE DÉBIT

Il importe maintenant de considérer un autre facteur de variation temporelle, le débit. En effet, on peut se demander (i) s'il y a une différence de débit entre le discours citant et le discours cité; (ii) si cette différence varie selon le type de discours rapporté; (iii) si le débit moyen de l'énoncé varie selon le type de discours rapporté; (iv) si l'exclusion de la durée des pauses silencieuses maintient la caractérisation des types de discours rapporté, le cas échéant.

Le débit est donc calculé dans le discours citant, dans le discours cité, puis dans l'ensemble de l'énoncé, et ce pour chaque type de discours rapporté. Il est ensuite traité de deux façons différentes : la vitesse de locution (VL) et la vitesse d'articulation (VA). La vitesse de locution tient

compte du temps de locution (TL), c'est-à-dire que sont inclus les temps de pause silencieuse. Ce qui donne la formule suivante :

$$VL = \frac{\text{nb. total de syllabes dans la séquence X 1000}}{\text{TL (ms)}} = \text{nb. de syllabes/seconde}$$

La vitesse d'articulation ne prend en compte que le temps d'articulation (TA), c'est-à-dire que sont exclus les temps de pause silencieuse, ce qui peut être ainsi formulé :

$$VA = \frac{\text{nb. total de syllabes dans la séquence X 1000}}{\text{TA (ms)}} = \text{nb. de syllabes/seconde}$$

Les pauses d'hésitation, de même que les répétitions, les troncations et les faux-départs sont considérés comme des syllabes et leur durée est toujours prise en considération. La méthode de calcul suivie est celle de GROSJEAN et DESCHAMPS (1972).

Dans les trois types de discours rapporté, la comparaison entre les débits en discours citant et ceux en discours cité est effectuée à l'aide d'un test *t* pairé. La comparaison du débit des discours citants en DD et en DI par rapport à celui en DD∅ est réalisée à l'aide d'un test *t* avec approximation de Satterthwaite.

La *Figure 3.7a* illustre les résultats pour la vitesse de locution du discours citant (VL1) par rapport à celle du discours cité (VL2). La *Figure 3.7b* présente les résultats pour la vitesse de locution de l'énoncé (VLÉ).

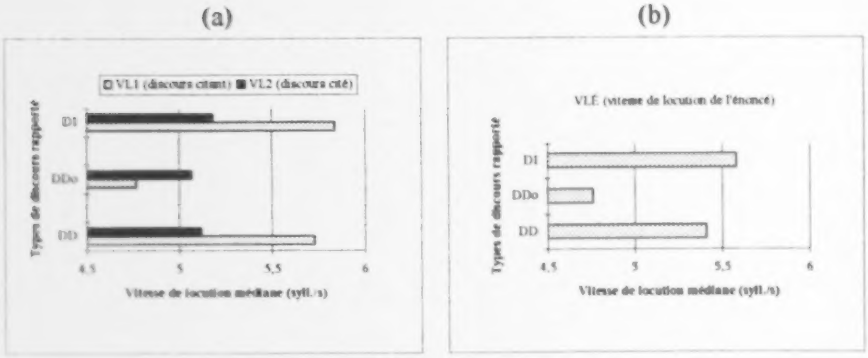


Figure 3.7 Le débit. (a) La vitesse de locution (VL) médiane (syll./s) en discours citant (VL1), en discours cité (VL2) et (b) dans l'ensemble de l'énoncé (VLÉ), dans les types de discours rapporté : discours direct avec verbe *dire* (DD), discours direct sans verbe *dire* (DDø) et discours indirect (DI).

Il se dégage de la *Figure 3.7a* que :

1) La vitesse de locution du discours citant (VL1) est significativement plus rapide que la vitesse de locution du discours cité (VL2) en discours direct avec verbe *dire* (DD; $p < 0,001$), comme en discours indirect (DI; $p < 0,05$), ce qui n'est pas le cas en discours direct sans verbe *dire* (DDø; n.s.).
2) La vitesse de locution du discours citant (VL1) en discours direct sans verbe *dire* (DDø) est significativement plus lente ($p < 0,001$) qu'en discours direct avec verbe *dire* (DD) et qu'en discours indirect (DI)[18].

La *Figure 3.7b* illustre d'autre part que :

3) la vitesse de locution de l'énoncé (VLÉ) en discours direct sans verbe

[18] La vitesse de locution du discours citant est semblable en DD et en DI, mais bien différente en DDø. Le test *t* avec approximation de Satterthwaite permet de réunir deux groupes dont le nombre de données n'est pas égal (ici DD et DI) et d'en faire un seul pour le comparer à un autre groupe (ici DDø).

dire (DDø) est aussi significativement plus lente ($p < 0,01$) qu'en discours direct avec verbe *dire* (DD) et qu'en discours indirect (DI).

On peut donc dire que la vitesse de locution du discours citant (VL1), comme la vitesse de locution de l'énoncé (VLÉ) du discours direct sans verbe *dire* (DDø), s'oppose à celle des deux autres types de discours rapporté.

Cependant, sachant que le discours direct sans verbe *dire* (DDø) est le type de discours rapporté qui présente le plus souvent une pause silencieuse à la transition, on peut supposer que la suppression de la durée de ces dernières dans le calcul du débit, ce qui est désigné comme étant la vitesse d'articulation, risque de neutraliser cette tendance. Si tel est le cas, les différences relevées dans la vitesse de locution ne sont pas véritablement significatives, mais sont plutôt la conséquence de la présence répétée d'une pause silencieuse. Dans la *Figure 3.8a*, la vitesse d'articulation médiane est donc présentée pour le discours citant (VA1) et pour le discours cité (VA2); la *Figure 3.8b* présente la vitesse d'articulation pour l'ensemble de l'énoncé (VAÉ). Les trois types de discours rapporté sont comparés.

(a) (b)

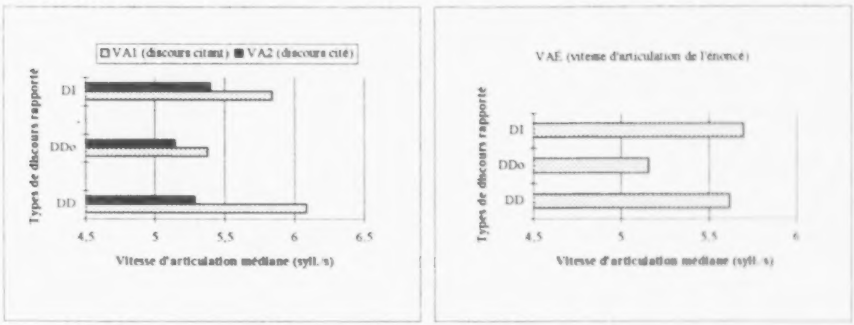


Figure 3.8 Le débit. (a) La vitesse d'articulation (VA) médiane (syll./s) en discours citant (VA1), en discours cité (VA2) et (b) dans l'ensemble de l'énoncé (VAÉ), dans les types de discours rapporté : discours direct avec verbe *dire* (DD), discours direct sans verbe *dire* (DDø) et discours indirect (DI).

Il apparaît dans la *Figure 3.8* (*a* et *b*) que l'ensemble des distinctions établies en considérant la vitesse de locution, entre discours direct sans verbe *dire* (DDø) d'une part et discours direct avec verbe *dire* (DD) et discours indirect (DI) d'autre part, sont maintenues en considérant la vitesse d'articulation, qui exclut pourtant le calcul des pauses silencieuses; ($p < 0,05$). La longueur des énoncés ne peut être évoquée puisque la médiane du nombre de syllabes par énoncé est tout à fait comparable d'un type de discours à l'autre : 13 syllabes en discours direct avec verbe *dire*, 14,5 syllabes en discours direct sans verbe *dire* et 13 syllabes en discours indirect.

Dans le cas présent, de tels résultats tendent à montrer qu'un ralentissement du débit peut suppléer à des «faiblesses» syntaxiques, en l'occurrence à l'absence du verbe *dire*.

D'un point de vue plus général, il est à souligner que ce maintien de la tendance en vitesse d'articulation confirme les résultats de MILLER *et al.* (1984:222). En effet, ceux-ci affirment, contrairement à GOLDMAN-EISLER (1968) pour l'anglais, à MALÉCOT *et al.* (1972) et à GROSJEAN et DESCHAMPS (1975:155) pour le français, que la vitesse d'articulation est plus variable qu'on ne l'a dit.

> It would appear that too much emphasis has been placed on the rather "low" variation in articulation rate as compared to the rather "high" variation in pause rate (the magnitude of variation issue) and this has led researchers to conclude that articulation rate per se does not vary substantially either between or within speakers. (MILLER *et al.* 1984:222)

Le débit ressort comme un paramètre prosodique utilisé non pas pour opposer les deux grands types de discours rapporté : le discours direct avec verbe *dire* (DD) présente des caractéristiques en tous points semblables à celle du discours indirect, mais plutôt pour suppléer à une «faiblesse» syntaxique : un ralentissement notable du débit en discours direct sans verbe *dire* (DDø) par rapport au débit du discours direct avec verbe *dire* (DD), que l'on considère la vitesse de locution (VL) ou la vitesse d'articulation (VA).

3.2.5 EN SOMME

De cette analyse de la production du discours rapporté dans une perspective syntaxique, il ressort que les tendances fréquencielles et temporelles distinctives des /discours directs/ et indirects sont plutôt relatives. L'examen approfondi de 215 énoncés de discours rapporté oral spontané, à partir de la fréquence fondamentale, de l'effet des «marqueurs transitionnels» sur celle-ci, des pauses et du débit, montre les caractéristiques prosodiques suivantes.

1) Le discours indirect

Le discours indirect est le type de discours rapporté le plus stable.

- 54% des *groupes intonatifs avant la transition* sont montants (M);
- 65% des transitions sont descendantes (D);
- 46% des profils d'énoncé sont descendants (D).
- L'ampleur des écarts médians reflète les tendances des fréquences relatives.
- Les patrons prosodiques significatifs pour l'énoncé de forme indirecte sont : MD-(D, P) ou PD-D (P = plat). Ils correspondent à 60% des énoncés.
- Il n'y a jamais de «marqueurs transitionnels», qu'il s'agisse d'un marqueur d'attaque (élément discursif, mais tout de même porteur de caractéristiques prosodiques), d'une pause silencieuse ou d'une pause d'hésitation.

2) Le /discours direct/

Les discours directs, avec ou sans verbe *dire* (/discours direct/), sont plus variables; ils se distinguent, en partie, du discours indirect par opposition. En effet, les traits tendanciels du discours indirect se trouvent en pourcentage beaucoup plus faible en /discours direct/.

- Seulement 21% des /discours directs/ ont des *groupes intonatifs avant la transition* montants; le *groupe intonatif avant la transition* est surtout plat (59%).

- 20% des /discours directs/ ont des transitions descendantes; la transition tendancielle est plate (40%) ou montante (40%). À la transition, l'ampleur des écarts médians est de 0,78 DT, donc plate.
- 21% des profils d'énoncé sont descendants; les profils sont plutôt montants (46%).
- Les patrons prosodiques significatifs pour l'énoncé rapporté directement sont : P (D)M-X. Ils correspondent à 34% des énoncés.
- 58% des /discours directs/ ont un «marqueur transitionnel».
- L'effet du «marqueur transitionnel» sur F_0 n'est pas significatif lorsqu'il s'agit d'un marqueur d'attaque, mais le devient lorsqu'il s'agit de la pause silencieuse (du moins du point de vue de la fréquence d'occurrence des tendances) : celle-ci diminue le nombre de transitions plates en DD (52% par rapport à 10%) et augmente le nombre de transitions montantes en DD∅ (40% par rapport à 58%).

Malgré un ensemble de tendances significatives, une marge de fluctuation prosodique relativement importante demeure dans la réalisation du /discours direct/.

3) Différences entre DD et DD∅

La distinction prosodique entre discours directs avec un verbe *dire* et discours directs sans verbe *dire* n'est pas très marquée. Rien de significatif du point de vue fréquenciel. Les caractéristiques distinctives du discours direct sans verbe *dire* tiennent plutôt :

- à une fréquence relative plus élevée de «marqueurs transitionnels» en DD∅ (70% par rapport à 46%);
- à un débit plus lent. Toutefois, cette caractéristique ne fait pas que l'opposer au discours direct avec verbe *dire*, mais aussi au discours indirect. En effet, discours direct avec verbe *dire* et discours indirect ont des caractéristiques de débit tout à fait semblables.

La relation entre prosodie et syntaxe s'avère donc relative. D'une part, la stabilité syntaxique du discours indirect, en fait il s'agit d'une complétive, se reflète, du moins partiellement, dans une certaine stabilité prosodique. D'autre part, en ce qui concerne le /discours direct/, les tendances prosodiques sont variables. Du point de vue fréquenciel, cette variabilité n'est pas le reflet de l'instabilité syntaxique du /discours direct/ : absence de verbe *dire* et/ou présence d'un «marqueur transitionnel». Mais alors, comment expliquer cette variabilité fréquencielle ? Simple produit du hasard ? Peu probable. Calque des attitudes ? Difficile à vérifier en raison du grand nombre d'attitudes et de la difficulté de définir chacune rigoureusement. Corrélat d'éléments discursifs ? Il s'agit là d'une piste vérifiable qui fait l'objet de la prochaine section.

3.3 PROSODIE ET DISCOURS

Si l'on admet avec GUAÏTELLA (1991:141) que la prosodie du spontané a plus à voir avec le sens qu'avec la syntaxe, la relation entre la prosodie du discours rapporté spontané et sa valeur discursive devrait être d'une qualité à tout le moins comparable à la relation entre la prosodie et la valeur syntaxique.

Contrairement à certains analystes du discours (dont AUTHIER:1978; MAINGUENEAU: 1981,1986, 1987; de GAULMYN:1992), VINCENT et DUBOIS (1996) ne se prononcent pas sur le marquage prosodique des différents types de discours rapporté. Cependant, elles proposent que les énoncés rapportés, directs et indirects, entretiennent des liens différents avec la reproduction de propos. Cette proposition permet de classer systématiquement les énoncés sur une base discursive plutôt que syntaxique et offre une voie de vérification privilégiée des rapports éventuels entre prosodie et discours. VINCENT et DUBOIS (1996:50-54) ont déterminé cinq types d'emploi du discours rapporté:

1) **La *reproduction***. Les énoncés contiennent tous les indices qu'un événement source a eu lieu : temps passé, énonciateur défini, destinataire défini et contexte défini. Ils sont associés au discours narratif. Ces énoncés constituent 39% des discours rapportés de leur corpus.

(8) Elle lui a dit : «Oui t'as été élevé comme un vrai fou.» (6229-79)

2) **La *pseudo-reproduction***. Les énoncés laissent croire que les paroles citées ont déjà été dites, temps passé et énonciateur défini, mais ni le destinataire, ni le contexte ne sont clairement définis (8% des énoncés).

(9) On a dit : «Coudon il-y-a quelque chose en quelque part.» (6189-79)

3) **L'*actualisation***. Les énoncés rapportés actualisés ne constituent pas à proprement parler un événement de communication unique, mais représentent plutôt un prototype de plusieurs événements similaires (34% des énoncés). Ils ne sont pas définis comme des énoncés narratifs.

(10) C'est gênant se faire dire : «Vous avez pris du poids.» (4043-15)

4) **L'*invention***. Les énoncés rapportés inventés sont des propos qui n'ont jamais explicitement été dits. Si l'énonciateur, le destinataire et le contexte sont habituellement définis, le temps de l'événement n'est ni au passé, ni au présent (13% des énoncés).

(11) Pour moi vous allez dire : «À soir elle les hommes là.» (4034-15)

5) **L'*assertion***. Cet emploi est le plus éloigné de ce qu'on appelle habituellement le discours rapporté. L'énonciateur est toujours le locuteur *je*, ou un énonciateur collectif incluant le *je*. Il n'y a aucun événement passé, aucun contexte défini. Si on enlevait le verbe de parole, l'énoncé pourrait s'intégrer au discours en cours. Cet emploi est plus fréquent en discours indirect qu'en discours direct (18% par rapport à 4%) et constitue 6% des emplois du discours rapporté.

(12) Je te dis pas qu'il me n'empruntait pas. Tsé euh. (6467i-59)

Le *Tableau 3.7* illustre les indices de reproduction de parole pour chacun des types d'emploi des énoncés rapportés.

Tableau 3.7 Indices de reproduction de paroles pour les types d'emploi discursif du discours rapporté (VINCENT et DUBOIS, 1996:54). R (*reproduction*); P-R (*pseudo-reproduction*); Act (*actualisation*); I (*invention*); Ass (*assertion*).

Emplois du discours rapporté	R	P-R	Act	I	Ass
Indices de reproduction					
Événement passé	+	+	+/-	-	-
Énonciateur défini	+	+	+/-	+	+**
Destinataire défini*	+	-	+/-	+	+/-
Contexte défini	+	-	+/-	+/-	-
Taux d'utilisation	39%	8%	34%	13%	6%

* Autre que réfléchi

** L'énonciateur ne peut être que *je*

La valeur de reproduction, de *(+) narrative* à *(-) narrative*, de ces emplois est ainsi représentée (VINCENT et DUBOIS, 1996:54).

reproduire pseudo-reproduire actualiser inventer // asserter

(+) narrative ——————————→ *(-) narrative*

Dans l'ensemble, la classification d'un énoncé dans un type d'emploi est assez nette. Il reste cependant que la détermination des indices de reproduction est parfois délicate. Ainsi dans l'exemple (13) où une adolescente rapporte des propos d'adultes, on peut hésiter entre l'*actualisation* ou l'*invention* :

(13) Comme le vin :«Ah celui-là il est bon celui-là il est bon.» (6143ø-131) Ben non.

Ces propos sont-il prononcés régulièrement devant elle ? Il s'agirait alors d'une *actualisation*. Ou sont-ils inventés par elle pour démontrer

l'attitude de adultes ? Il s'agirait alors d'une *invention*. C'est le choix qui a été fait ici en raison de l'absence de tout modalisateur (du type *des fois*, par exemple).

Ceci dit, la classification est plutôt efficace. Certes, le critère de la valeur narrative ne recouvre pas entièrement ce que l'on a nommé sous le terme générique de valeur discursive. Par exemple, il est raisonnable de penser que la valeur d'attitude évoquée par AUTHIER (1978) et par MAINGUENEAU (1986) peut fort bien entraîner des prosodies différentes pour des énoncés qui appartiennent pourtant à un même type d'emploi, comme le type d'emploi *reproduction* dans les exemples (14) et (15). Dans l'exemple (14), l'attitude du rapporteur par rapport au paroles citées apparaît négative :

(14) C'est une fille qui était pas rodée pi tout 'suite # :«Ces deux-là dehors.» (4065⌀-15)

alors que dans l'exemple (15), elle paraît plus positive.

(15) Les autres m'encourageaient :«Cours cours Bina.» (5609⌀-108)

Il est probable que ces attitudes différentes soient aussi des facteurs de variation prosodique.

En ce sens, les indices de reproduction de paroles ne constituent pas un moyen de vérification parfaitement satisfaisant des relations entre la prosodie du discours rapporté et sa valeur discursive. Toutefois, ils n'en constituent pas moins un moyen approprié pour observer une éventuelle relation entre la prosodie et la valeur discursive.

Dans le *Tableau 3.8*, on retrouve les 163 énoncés en /discours direct/ et les 52 énoncés en discours indirect, classés selon les 5 catégories d'emploi, telles que décrites dans le *Tableau 3.7*.

84

Tableau 3.8 Distribution des énoncés rapportés selon leur valeur narrative (corpus à l'étude). R (*reproduction*); P-R (*pseudo-reproduction*); Act (*actualisation*); I (*invention*); Ass (*assertion*).

n = 215	R N	%	P.-R N	%	Act N	%	I N	%	Ass N	%
/DD/(n=163)	41	25%	14	9%	68	42%	33	20%	7	4%
DI (n=52)	23	44%	1	2%	10	19%	1	0%	17	33%
TOTAL (n=215)	64	30%	15	7%	78	36%	34	16%	24	11%

Il est à remarquer que les énoncés du présent corpus n'ont pas été sélectionnés en fonction de leur valeur discursive, mais plutôt de leur structure syntaxique (DD, DD⌀ et DI). La répartition /discours direct/ et discours indirect dans le type d'emploi aux extrêmes du continuum (*reproduction* et *assertion*) n'est pas proportionnelle à celle observée dans le corpus des 4860 occurrences de VINCENT et DUBOIS. En effet, seulement 25% des /discours directs/ du présent corpus appartiennent au type d'emploi *reproduction* par rapport à 40% dans le corpus de VINCENT et DUBOIS alors que 33% des discours indirects appartiennent au type d'emploi *assertion* par rapport à 18% dans le corpus de VINCENT et DUBOIS. Si la valeur discursive a un effet sur la prosodie du discours rapporté, on peut penser qu'un corpus plus représentatif des types d'emploi discursif aurait donné des tendances plus fortes que celles observées à travers la classification syntaxique. C'est ce qu'il convient maintenant d'examiner.

Hypothèses

À partir de ce qui a été observé dans l'analyse de la relation entre la prosodie et la syntaxe, les hypothèses suivantes sont posées:

Pour la mélodie :

1. Compte tenu que la mélodie du discours indirect est relativement stable, le plus souvent conforme à la mélodie de l'énoncé déclaratif,
l'hypothèse suivante est posée : la mélodie du discours indirect ne serait pas vraiment influencée par la valeur discursive.

2. Compte tenu que la mélodie du discours direct est variable et que les seules différences syntaxiques (occasionnellement, absence du verbe *dire* et/ou présence de marqueurs d'attaque) ne suffisent pas à justifier cette variabilité,
l'hypothèse suivante est posée : la mélodie du discours direct serait influencée par la valeur discursive.

3. Compte tenu qu'en discours direct spontané, les tendances observées pour la mélodie sont un *groupe intonatif avant la transition* plat, une transition aussi souvent plate que montante et un profil global de l'énoncé montant et compte tenu qu'en discours direct lu, c'est la transition montante qui est caractéristique du discours direct (LÉANDRI:1993),
l'hypothèse suivante est posée : plus la valeur narrative de l'énoncé rapporté en discours direct serait forte, plus les tendances observées pour F_0 seraient présentes. La transition montante serait davantage associée à une valeur narrative forte que la transition plate.

Pour la pause :

Compte tenu que la pause à la transition est fortement reliée au plan syntaxique : (i) occasionnellement présente à la transition en discours direct, toujours absente en discours indirect, (ii) plus fréquente lorsque le verbe introducteur est absent (iii) plus longue à la transition qu'ailleurs dans l'énoncé,
l'hypothèse suivante est posée : la pause ne serait pas particulièrement sensible à la valeur discursive.

SS

Pour le débit :

Compte tenu qu'en discours spontané, il a été observé que le débit (nombre de mots par minute) est plus rapide en discours narratif (LAFOREST 1995)[19] et que la *reproduction* est un type d'emploi associé au discours narratif,
l'hypothèse suivante est posée : le débit moyen de l'énoncé de type *reproduction* serait plus élevé que celui des autres types d'emploi.

3.3.1 LA FRÉQUENCE FONDAMENTALE et la valeur narrative

En discours direct

D'une part, dans l'ensemble, les caractéristiques fréquencielles sont ambivalentes lorsqu'on les analyse du seul point de vue de la structure syntaxique non subordonnante du /discours direct/ : la proportion de transitions plates (40%) est aussi importante que celle des transitions montantes (40%). D'autre part, le *groupe intonatif avant la transition* est plus souvent plat (59%) et le profil plus souvent montant (46%).

Les résultats pour la fréquence fondamentale sont illustrés par des figures qui comparent la fréquence relative tendancielle des configurations tonales obtenue à partir du regroupement syntaxique (/DD/) et la fréquence relative des configurations tonales obtenue à partir du regroupement discursif, de la valeur narrative la plus forte à la valeur narrative la moins forte (i.e. de la *reproduction* à l'*assertion*).

---

[19] Il faut cependant préciser que LAFOREST a analysé le débit de récits, genre discursif complexe qu'il ne faut pas confondre avec la fonction narrative de certains énoncés.

Les résultats sont validés par un test khi deux de tendance[20].

(i) Le groupe intonatif avant la transition

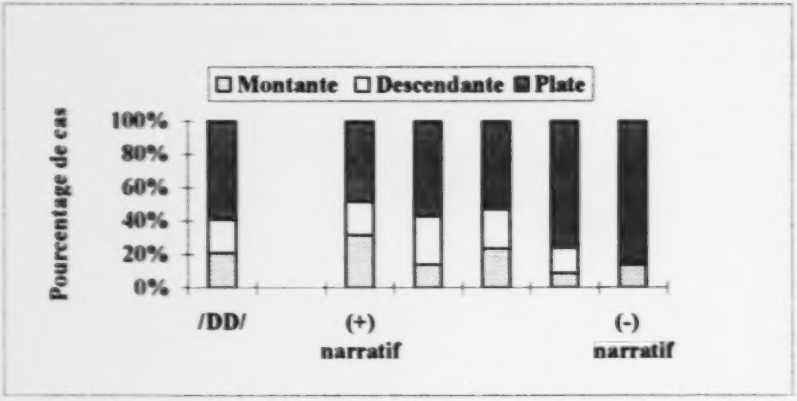


Figure 3.9 *Groupe intonatif avant la transition*. Fréquence relative des configurations tonales pour l'ensemble des discours direct (/DD/) et pour chacun des cinq emplois discursifs.

On voit dans la *Figure 3.9* que pour le *groupe intonatif avant la transition*, de façon globale, le pourcentage de configurations tonales plates tend à augmenter légèrement au fur et à mesure que la valeur narrative décroît; ($p < 0,05$). Cependant, les écarts médians du *groupe intonatif avant la transition* varient peu d'un type d'emploi à l'autre (de 0 à 0,45 DT). Il reste qu'en fréquence relative la tendance plate de ce domaine est inversement proportionnelle à la valeur narrative : plus la valeur narrative croît, plus le nombre de *groupes intonatifs avant la transition* plats diminue. L'hypothèse de la

[20] Un test de khi deux de Pearson pourrait vérifier si la proportion de *groupes intonatifs avant la transition*, par exemple, varie selon le type d'emploi discursif. Un test de khi deux de tendance est un peu plus précis : il vérifie si la proportion de *groupes intonatifs avant la transition* augmente, ou diminue, selon le type d'emploi discursif.

SS

corrélation entre la valeur narrative forte et les tendances dégagées par le regroupement syntaxique est ici infirmée.

(ii) La transition

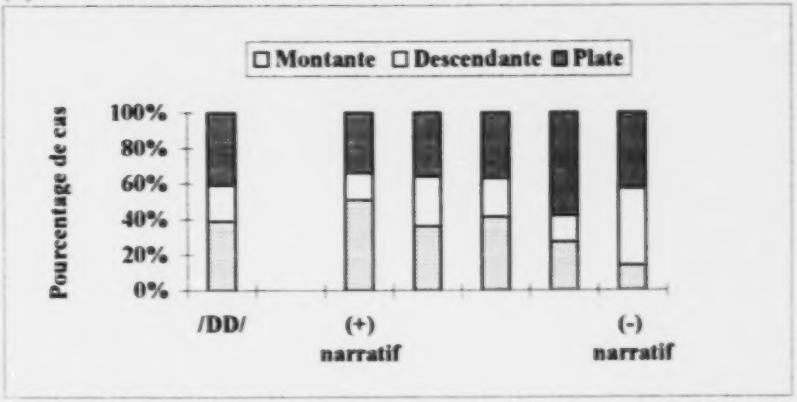


Figure 3.10 Transition. Fréquence relative des configurations tonales pour l'ensemble des discours directs (/DD/) et pour chacun des cinq emplois discursifs.

La *Figure 3.10* illustre d'une part que le pourcentage de transitions plates n'est pas corrélé à la valeur narrative; d'autre part, plus la valeur narrative est forte, plus le pourcentage de transitions montantes augmente (de 51% en *(+) narratif* à 14% en *(-) narratif*); ($p < 0,05$). Parallèlement, les écarts médians de la transition suivent une échelle décroissante tout à fait représentative de la valeur narrative, comme l'illustre le *Tableau 3.9*.

Tableau 3.9 Transition. Écarts médians (en demi-tons, DT) selon les cinq types d'emploi discursif en /discours direct/.

Types d'emploi discursif	Transition
Reproduction	2,37 DT
Pseudo-reproduction	1,30 DT
Actualisation	0,79 DT
Invention	0,00 DT
Assertion	-1,38 DT

L'ambivalence des valeurs fréquencielles laissée par la mise en rapport entre les caractéristiques fréquencielles et la syntaxe est ici expliquée. La configuration de F_0 est tendanciellement reliée à la valeur discursive de l'énoncé rapporté : quand la valeur narrative est forte, la transition a tendance à être montante. D'autre part, si en termes de fréquence d'occurrences, la transition plate n'est pas reliée à la valeur discursive du discours rapporté, les écarts médians s'apparient à une transition plate dans les types d'emploi discursif autres que la *reproduction* (de 1,30 DT à -1,38 DT). Pour ce domaine stratégique qu'est la transition, l'hypothèse est confirmée.

Le domaine du profil global de l'énoncé n'est pas significativement sensible à la valeur narrative de l'énoncé rapporté.

En discours indirect

En ce qui concerne le rapport entre la structure syntaxique du discours indirect et les caractéristiques fréquencielles, on se rappelle qu'il est relativement stable : 54% des *groupes intonatifs avant la transition* sont montants, 65% des transitions sont descendantes alors que 46% des profils sont descendants. Il est tout de même opportun d'examiner si cette stabilité se maintient lorsque les énoncés de forme indirecte sont distribués en fonction de leur valeur narrative. Étant

200

donné le peu de cas représentatif de la *pseudo-reproduction* (1) et de l'*invention* (1), ces types d'emploi n'ont pas été retenus.

Lorsque les énoncés de forme indirecte sont distribués en fonction de leur valeur narrative, même si on remarque que la transition descendante et le profil descendant augmentent dans l'*assertion*, aucune tendance statistiquement significative n'a été dégagée. Dans ce cas, prosodie et syntaxe sont corrélées, autant que faire se peut en oral spontané.

En revanche, l'apparente ambivalence des tendances fréquencielles à la transition des /discours directs/ paraît associable à la valeur discursive de l'énoncé. Il y a là matière à supposer qu'une même intonation entretient tantôt des liens avec la structure syntaxique, tantôt avec la valeur discursive.

3.3.2 LA PAUSE et la valeur narrative

Tel qu'attendu, il n'apparaît pas globalement y avoir de lien entre la fréquence d'apparition des pauses à la transition ainsi que leur durée et la valeur narrative de l'énoncé rapporté. On remarque toutefois que la valeur narrative la plus faible, l'*assertion*, n'a ni pause silencieuse, ni pause d'hésitation à la transition, comme en discours indirect. Il ne faut cependant pas perdre de vue que le nombre de pauses déjà limité (29/163 énoncés) devient minime lorsqu'il est fractionné selon les cinq types d'emploi discursif. La question du débit est maintenant examinée.

3.3.3 LE DÉBIT et la valeur narrative

Le débit est traité en considérant la vitesse d'articulation (voir la section 3.2.4). Cette vitesse d'articulation est examinée dans les cinq

types d'emploi discursif pour le /discours direct/[21] et dans seulement trois types d'emploi pour le discours indirect (*reproduction, actualisation* et *assertion*; les types *pseudo-reproduction* et *invention* ne sont représentés que par un énoncé chacun).

Des tests paramétriques (basés sur des moyennes) ont d'abord été effectués. Cependant, puisque le nombre d'énoncés dans certains types d'emploi discursif n'était pas toujours très élevé en /discours direct/ (n= 7, 14, 10, 17, 23) et que les distributions n'étaient pas toujours normales, des tests non paramétriques (basés sur des médianes) ont également été utilisés afin de valider les conclusions des tests paramétriques. Les deux types de test donnent des résultats semblables. Ce sont les tests basés sur les moyennes qui sont présentés ci-après. Il s'agit principalement d'une analyse de la variance de mesures répétées.

En discours direct

La première question à se poser : est-ce qu'il y a une interaction significative entre les différences VA1 (vitesse d'articulation en discours citant) et VA2 (vitesse d'articulation en discours cité) et le type d'emploi discursif ? Autrement dit, est-ce que les différences entre VA1 et VA2 varient selon le type d'emploi discursif ? La différence n'est pas significative. Elle le devient cependant pour chacun des types d'emploi pris séparément. Ce fait n'ajoute rien à la relation entre le débit et la valeur discursive, mais il tend à montrer que la rupture de débit entre le discours citant et le discours cité est un indice de discours rapporté directement. En revanche, le débit moyen des énoncés de type *reproduction* semble plus élevé que celui des autres types de discours rapporté ($p < 0,01$), tel que montré dans la *Figure 3.11*.

---

[21] Étant donné les caractéristiques spécifiques du débit en DDe, il aurait été intéressant de prendre en compte les deux sous-catégories de discours direct. Malheureusement, le morcellement occasionné par la subdivision en cinq types d'emploi discursif ne le permettait pas.

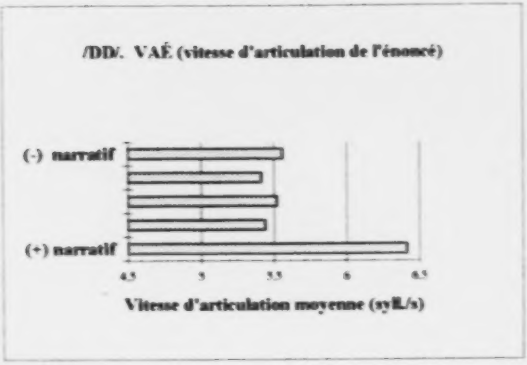


Figure 3.11 Le débit (/DD/). La vitesse d'articulation moyenne dans l'ensemble de l'énoncé (VAÉ) pour les cinq types d'emploi discursif.

À l'aide de la méthode des comparaisons multiples de Tukey, il peut être précisé que la vitesse d'articulation en *reproduction* est statistiquement différente de la vitesse d'articulation en *actualisation* et en *invention*. Les deux autres types ne sont pas montrés significativement différents du type *reproduction* en raison du petit nombre d'énoncés qu'ils contiennent (14 en *pseudo-reproduction* et 7 en *assertion*).

En discours indirect

Sans être tout à fait significative, il apparaît dans le *Tableau 3.10* que l'interaction entre la différence VA1/VA2 et le type d'emploi discursif est possible, en tout cas plus qu'en /discours direct/.

Tableau 3.10 Le débit (DI). Différence de vitesse d'articulation (syll./s) entre le discours citant (VA1) et le discours cité (VA2) selon les trois types d'emploi discursif.

Types d'emploi discursif	VA1 moyenne - VA2 moyenne en DI
Reproduction	1,18 syll./s
Actualisation	0,21 syll./s
Assertion	0,11 syll./s

En effet, il apparaît que la différence VA1/VA2 est plus importante en *reproduction* qu'en *actualisation* et en *assertion* (1,18 par rapport à 0,21 et 0,11 syll./s), bien que la valeur de p soit faible ($p = 0,0967$).

Par ailleurs, l'effet simple observé en discours direct à propos de la vitesse d'articulation moyenne (VAÉ) ne se manifeste pas en discours indirect. VAÉ n'est pas significativement plus rapide en *reproduction* que dans les deux autres types d'emploi discursif

De cette analyse du débit, il ressort que la corrélation entre les valeurs narratives et le débit est relativement marquée :

1) La différence de débit entre discours citant (VA1) et discours cité (VA2) ne semble pas très liée au type d'emploi discursif : pas du tout en /discours direct/ et légèrement en discours indirect.
2) En revanche, en /discours direct/, la vitesse d'articulation de l'énoncé (VAÉ) est plus rapide dans l'emploi à haute valeur narrative (la *reproduction*).

3.3.4 EN SOMME

L'analyse de la production du discours rapporté d'un point de vue discursif, particulièrement du /discours direct/, montre que des tendances fréquencielles et temporelles sont corrélées à la valeur discursive de l'énoncé. Du point de vue temporel, seul le débit est mis en cause; en effet, la fréquence d'occurrence des pauses ainsi que leur durée n'apparaissent pas reliées aux valeurs discursives.

1) Le discours indirect

On se rappelle que la corrélation entre la structure syntaxique, relativement stable du discours indirect, et les marques prosodiques s'est avérée assez régulière. La forme indirecte de l'énoncé rapporté est peu touchée par la valeur discursive.

- La valeur narrative n'influence pas véritablement les tendances fréquencielles du discours indirect.
- Les différences de débit entre VA1 et VA2 des discours indirects sont plus importantes lorsque la valeur narrative est forte (cf. *Tableau 3.10*), mais la valeur p est plutôt faible (p = 0,0967).

2) Le /discours direct/

Dans ce cas, le regroupement discursif est révélateur de la corrélation des marques prosodiques, aussi bien fréquencielles que temporelles, avec la valeur narrative de l'énoncé.

- Plus la valeur narrative est forte, moins le *groupe intonatif avant la transition* présente une configuration tonale plate (cf. *Figure 3.9*) et plus la transition présente une configuration tonale montante (cf. *Figure 3.10*).
- Le débit (VAÉ) est significativement plus rapide lorsque la valeur narrative est à son plus haut degré (cf. *Figure 3.11*).

Dans le cas du discours rapporté, l'analyse de la relation entre prosodie et discours s'avère donc un complément nécessaire à l'analyse de la relation entre prosodie et syntaxe, particulièrement en /discours

direct/. En effet, elle permet de préciser que le marquage intonatif par une transition montante ainsi qu'une augmentation du débit moyen de l'énoncé sont des caractéristiques prosodiques corrélées à une valeur narrative forte (le type d'emploi de la *reproduction*).

3.4 DISCUSSION

L'analyse de la production du présent corpus de discours rapportés a montré que prosodie, syntaxe et discours sont interreliés, dans une certaine mesure. Affirmées par les analystes du discours, confirmées par l'analyse prosodique d'un corpus lu, la présente analyse prosodique d'un corpus spontané présente des résultats plus nuancés. En effet, les relations entre prosodie et syntaxe, sont d'une part, relatives, et d'autre part, plus marquées en discours indirect qu'en discours direct. Quant aux relations entre prosodie et valeur discursive, elles n'avaient, à ma connaissance, fait l'objet d'aucune étude quantitative auparavant. En discours direct, elles paraissent ici tendanciellement fonction de la valeur narrative.

Les fluctuations entre prosodie, syntaxe et discours, qui viennent d'être observées au plan de la production, sont également visibles dans la transcription prosodique. De plus, elles se confirment dans des exemples d'énoncés rapportés tirés du corpus d'entrevues télévisées de MERTENS (1987).

3.4.1 LA REPRÉSENTATION DE L'INTONATION ET L'ASSOCIATION PROSODIE, SYNTAXE ET DISCOURS

La représentation de l'intonation ici utilisée, basée principalement sur les données acoustiques transcrites de façon abstraite à l'aide de quatre tons (l'infra-bas, le suraigu, le haut et le bas) puis regroupées selon la règle de hiérarchie de MERTENS (1987), devrait fournir, dans la mesure du possible, une représentation assez

objective des regroupements intonatifs. Cette représentation permet d'observer, d'une autre façon, un certain nombre de cas qui contreviennent à l'association prosodie, syntaxe et discours.

Les groupes intonatifs sont identifiés par les parenthèses ( ), les crochets [ ] regroupent les groupes intonatifs en unités sémantiques selon le type d'accent final (règle de hiérarchie, MERTENS 1987).

Par exemple, chaque sous-catégorie de discours direct présente des cas où discours citant et discours cité appartiennent au même groupe intonatif[22] (en gras dans les exemples) alors que du point de vue syntaxique, ils constituent des propositions différentes; du point de vue discursif, le cas se produit aussi bien dans des énoncés à forte valeur narrative, exemple (16), où normalement discours citant et discours cité devraient être plus distincts, que dans les énoncés à faible valeur narrative.

(16)	**[(Elle dit : «Vous êtes ben**					**)] [(vous autres.»)]**		
	(*reproduction*) (4664-59)							
ms	93	115	138	317		205		
Hz	172	143	170	*(162-227)*208		204		
DT		-3,2	2,99	*(5,84)* 3,49				
TP	[(bb	\bb	bb	BH	)] [(	HH	)]	

D'autre part, il arrive occasionnellement (21/163), particulièrement dans les /discours directs/, que ceux-ci sont porteurs d'une prosodie *typique* —dans le sens des tendances relevées— du discours indirect : montée dans la partie citante, et/ou chute à la transition entre discours citant et discours cité, et/ou profil global descendant, et/ou absence de «marqueur transitionnel», comme l'illustre l'exemple (17) On appellera ce

[22] On en retrouve également dans le sous-corpus de MERTENS dont il sera question en 3.4.2.

phénomène, un phénomène de «bascule»[23].

(17) [(Mais comme on dit toujours)] [(: «Il-y-a pas de sot métier.»)] (*actualisation*) (6176-76)

Hz	**104**	122	125	0	95	**125**		**83**	123	102	70	**82**
DT*[24]						4,75		-7,09				-4,11
												(104 à 82 Hz)

Ainsi dans ce discours direct avec le verbe *dire*, on trouve une montée dans la partie citante (4,75 DT), une chute à la transition (-7,09 DT), un profil global descendant (-4,11 DT) ainsi que l'absence d'un «marqueur transitionnel», toutes tendances caractéristiques du discours indirect. Dans les discours directs sans verbe *dire*, aucun énoncé ne présente toutes les caractéristiques *typiques* du type de discours opposé.

En discours indirect, un seul cas, l'exemple (18), présente une «bascule». La «creaky voice» (*) des deux dernières syllabes assurent d'un profil descendant. Seule la montée à la transition peut surprendre, et encore. Le /i/ de *dis* est sourd, et l'écart de la transition (3,61 DT) n'est donc plus intersyllabique.

(18) [(Fait-que)] [(je me dis)(que j'ai des chances.)] (*assertion*) (6705i-125)

Hz	**189**	227	**194**	0	**239**	*	*	
DT*					?3,61			

La stabilité prosodique du discours indirect s'oppose à nouveau à la variabilité prosodique du /discours direct/.

---

[23] Au moins deux des caractéristiques doivent être présentes pour que l'énoncé soit considéré comme prosodiquement «basculé».

[24] **DT*** indique que les écarts sont calculés à l'intérieur d'un domaine plutôt que d'une syllabe à l'autre : pour la partie citante, et/ou la partie citée, et/ou le profil global. Le calcul à la transition demeure intersyllabique, à moins qu'il ne n'agisse d'un marqueur polysyllabique. Dans ce cas, le calcul est fait de la dernière syllabe du discours citant à la dernière syllabe du marqueur, ou à la première qui suit le marqueur.

Quelques exemples tirés du corpus d'oral spontané de MERTENS (1987) vont dans le même sens de non correspondance catégorique entre prosodie, syntaxe et discours, particulièrement en /discours direct/.

3.4.2 COMPARAISON AVEC DES DISCOURS RAPPORTÉS TIRÉS DU CORPUS DE MERTENS (1987)

Il est intéressant de comparer les résultats ici obtenus avec ceux des énoncés en discours rapporté relevés dans le corpus de français hexagonal spontané de MERTENS (1987) : onze (11) énoncés en /discours direct/, dont un (1) seulement sans verbe *dire*, et six (6) énoncés en discours indirect. La distinction entre discours direct avec verbe *dire* et discours direct sans verbe *dire* n'a de ce fait pas été maintenue; les tendances des deux sous-catégories ont été réunies (/discours direct/). À partir de la transcription abstraite de MERTENS, les tendances fréquencielles montantes, descendantes ou plates des domaines significatifs dans la présente étude ont été relevées. Le *Tableau 3.11* présente la comparaison, du point de vue syntaxique.

Tableau 3.11 Structure syntaxique. Organisation prosodique tendancielle des /discours directs/ (/DD/) et des discours indirects (DI) relevés dans le corpus de MERTENS (1987), **en gras**, comparée à celle du présent corpus.

	/DD/		DI	
	MERTENS (n= 11)	Corpus à l'étude (n = 163)	MERTENS (n =6)	Corpus à l'étude (n =52)
GI AVANT TRANSITION				
• montant	**55%**	22%	**100%**	54%
TRANSITION				
• descendante	**45%**	20%	**100%**	65%
PROFIL DE L'ÉNONCÉ				
• montant	**67%**	46%		
• descendant			**67%**	46%

La première remarque concerne les taux toujours plus élevés dans le corpus de MERTENS. Certes le nombre d'énoncés comparés ainsi que le nombre de locuteurs comparés sont loin d'être proportionnels. De plus, le type d'oral est différent et la variété dialectale également. En effet, ce sont 11 /discours directs/ qui sont comparés à 163 et 6 discours indirects à 52; les discours rapportés tirés du corpus de MERTENS ne sont issus que de 3 locuteurs, par rapport à 16 dans le présent corpus, et ces locuteurs sont des «professionnels de la parole» (Françoise Giroud, Jacques Chirac et Jacques Chancel), tandis que les locuteurs du corpus à l'étude représentent divers statuts socioprofessionnels, mais aucun n'est un «professionnel de la parole». Il s'agit en outre d'entretiens télévisés, en français hexagonal, ce qui est bien différent des conversations courantes du présent corpus de français québécois. On peut ajouter à cela que 8 des 11 énoncés en discours direct sont du type d'emploi discursif de la *reproduction* (73%) par rapport à 25% dans le présent corpus; on se souvient que ce type d'emploi est plus souvent porteur des caractéristiques prosodiques tendancielles.

Malgré tout, dans les deux corpus, on constate que les grandes tendances sont les mêmes :

1) le discours indirect est plus stable que le discours direct;
2) la tendance montante de la partie citante est caractéristique du discours indirect;
3) la tendance descendante de la transition est également caractéristique du discours indirect;
4) la tendance descendante de l'énoncé est aussi caractéristique du discours indirect alors que la tendance montante est caractéristique du /discours direct/.

Cette comparaison, bien qu'elle porte sur un nombre limité de cas, tend à confirmer qu'en spontané, s'il existe une tendance globale à distinguer prosodiquement les deux grandes catégories de discours rapporté de la façon décrite ci-haut, il y a place pour des stratégies éminemment variables, particulièrement en ce qui concerne le /discours direct/.

210

En ce qui concerne la comparaison entre valeurs fréquencielles et valeurs discursives, le peu d'énoncés rapportés relevés dans le corpus de MERTENS rend difficile une véritable comparaison. En /discours direct/, aucune comparaison n'est possible puisque le corpus de MERTENS ne présente pas d'exemple à valeur narrative faible. D'autre part, en discours indirect, dans le corpus de MERTENS, encore plus que dans le corpus à l'étude, la valeur discursive n'a pas d'effet important sur la prosodie : les tendances fréquencielles de la *reproduction* (3 énoncés) sont tout à fait comparables à celles de l'*invention* et de l'*assertion* réunies (3 énoncés).

3.5 PROSODIE ET FACTEURS SOCIAUX

Il sera d'abord fait mention de quelques études socioprosodiques qui se sont intéressées aux relations entre la prosodie et l'âge, le sexe ainsi que l'appartenance socioprofessionnelle, les trois paramètres sociaux ici retenus. En tenant compte de ces facteurs, il y a répartition des locuteurs ainsi que distribution des énoncés selon les types de classification, syntaxique et discursive, utilisés précédemment. Les résultats sont présentés à partir de ces deux types de classification pour la configuration de F_0 à la transition, le débit et la pause silencieuse.

En raison du fractionnement des données occasionné par les précédentes subdivisions, particulièrement en discours indirect, l'étude contrastive entre les deux grands types de discours rapporté devient quelque peu délicate. L'analyse statistique des variables continues est faite à l'aide du test Mann-Whitney non pairé. Cette analyse n'est cependant menée qu'à titre exploratoire.

3.5.1 QUELQUES ÉTUDES SOCIOPROSODIQUES

Depuis le début des années soixante, bon nombre d'études se sont intéressées aux relations entre le langage et les facteurs sociaux. Les travaux en linguistique de LABOV (1964), de même que ceux en

111

ethnographie de la communication de GUMPERZ et HYMES (1964), sont pour ainsi dire à l'origine de la prise en compte systématique des caractéristiques sociales des locuteurs dans la description du langage. Toutefois, la plupart des travaux à perspective sociolinguistique s'intéressent davantage à la relation entre les caractéristiques sociales et les caractéristiques lexicales, grammaticales, phonétiques ou phonologiques qu'à la relation entre les caractéristiques sociales et les caractéristiques prosodiques. Les premières tentatives pour relier paramètres sociaux et prosodie relèvent autant des anthropologues que des linguistes. On y retrouve peu de données significatives : pas de théorie commune, pas de consensus sur la terminologie, mais l'orientation est donnée.

La collection d'articles réunis dans GUMPERZ et HYMES (1964) est à ce titre tout à fait exceptionnelle. Une décennie plus tard, CRYSTAL (1975:84-88) fait la revue d'une série d'études prosodiques effectuées en rapport avec des facteurs sociaux. La relation naturelle, universelle, entre les traits prosodiques et des paramètres sociaux n'est plus à démontrer, mais la progression dans la production d'études systématiques a tout de même été lente. On pourrait pointer la cause d'un développement tardif dans ce secteur en paraphrasant GOFFMAN qui qualifiait les aspects expressifs du discours de «greasy parts of speech» constituant de ce fait «The Neglected Situation» (GOFFMAN 1964:133). Il reste qu'au cours des dernières décennies, de plus en plus de chercheurs se sont intéressés, de façon méthodique, aux effets des facteurs sociaux sur la prosodie. Quelques travaux mettant en relation les facteurs sociaux retenus dans la présente étude avec l'intonation, la pause silencieuse et le débit sont signalés.

Ces travaux ne portent pas sur le discours rapporté (à ma connaissance, il n'en existe pas), mais il sera utile d'observer :

1) dans quelle mesure ces études sont comparables entre elles puis à la présente étude;
2) dans quelle mesure les corrélations observées dans ces études se retrouvent en discours rapporté;
3) dans quelle mesure ces caractéristiques prosodiques sont distinctives des types de discours rapporté.

Le *Tableau 3.12* présente des corrélations entre des facteurs sociaux et des facteurs prosodiques : F_0 et sexe (KEY:1972; BREND:1972; LAKOFF:1975; MCCONNELL-GINET:1978; MORA:1990), F_0, sexe, âge et statut socioprofessionnel (GUY *et al.*:1986), F_0, débit et sexe (RYALLS *et al.*:1994), débit et âge (RYAN:1972; RAMIG:1983; van RIE *et al.*:1995), débit, pause et sexe (BYRD:1994)[25].

[25] Chacune de ces études est présentée dans DEMERS 1996:265-271.

Tableau 3.12 Corrélations entre facteurs sociaux et facteurs prosodiques dans quelques études socioprosodiques.

SEXE	ÂGE	S. SOCIOPROFESSIONNEL
INTONATION, Femmes (F) 1) Terminal déclaratif montant • hésitation (Key:1972) • politesse (Brend:1972) • demande d'assentiment (Lakoff:1975) • vérification de la compréhension (Guy *et al.*:1986) 2) Intonation plus «expressive» (Brend:1972; McConnell-Ginet:1978; Ryalls *et al.*:1994)	**Jeunes** •vérification de la compréhension (Guy *et al.*:1986)	**Défavorisés** Terminal déclaratif montant •demande d'assentiment (*Mora:1990) •vérification de la compréhension (Guy *et al.*:1986) **Favorisés** Intonation plus modulée (*Mora:1990)
DÉBIT, Hommes (H) 1) Plus rapide (*Mora:1990; Byrd:1994) 2) Pas de relation (Ryalls *et al.*:1994)	**Aînés** 1) Plus lent (Ryan:1972; Ramig:1983) 2) Plus lent chez les H; variable chez les F. (*van Rie *et al.*:1995)	
PAUSE Pas de relation (Byrd:1994)		

*Corpus d'oral spontané

Dans le *Tableau 3.12*, on décèle un certain nombre de difficultés liées à la mise en rapport de facteurs sociaux et de facteurs prosodiques. En effet, il y a chez les chercheurs une différence importante dans la façon d'interpréter la valeur sémantique d'un patron intonatif : un même patron intonatif, une montée terminale dans un énoncé déclaratif, est associé à l'hésitation (KEY:1972), à la politesse (BREND:1972), à une demande d'assentiment (LAKOFF:1975), toutes des interprétations plus liées à la soumission que celle de GUY *et al.* (1986) qui l'associent plutôt à la vérification de la compréhension de l'auditeur. La connotation de servilité se transforme en une connotation qu'on pourrait associer, par exemple, à un plus grand souci de communication. En plus de ces variations terminologiques, différents locuteurs peuvent utiliser un même contour intonatif avec une intention différente, ou inversement des contours intonatifs différents avec une même intention (BOLINGER:1964). Il n'est pas facile de s'y retrouver !

De plus, les échantillonnages issus de situations différentes, des corpus fabriqués (lus ou joués) par rapport à des corpus d'oral spontané, rendent périlleuse la comparaison des résultats. Par exemple, les études de BREND (1972) et de RYALLS *et al.* (1994), faites à partir de corpus fabriqués, montrent que l'intonation est plus «expressive» chez les femmes alors que l'étude de MORA (1990), faite à partir d'un corpus d'oral spontané, montre que l'intonation est plus modulée dans le groupe favorisé, hommes ou femmes. En revanche, l'étude de MORA (1990), faite à partir d'un corpus d'oral spontané, et l'étude de BYRD (1994), faite à partir d'un corpus lu, montrent que le débit est plus rapide chez les hommes que chez les femmes alors que l'étude de RYALLS *et al.* (1994), faite également à partir d'un corpus lu, montre qu'il n'y a pas de lien entre le débit et le sexe. BYRD (1994) tend à croire que la différence est due à la longueur des phrases : 13 syllabes dans son corpus et 5 à 7 syllabes dans le corpus de RYALLS *et al.* (1994). MORA (1990) ne fait pas mention du nombre de syllabes, mais de la durée des énoncés (entre 61 cs et 199 cs), BYRD (1994) fait état d'énoncés moyens de 265 cs chez les hommes et de 285 cs chez les femmes; RYALLS *et al.* (1994) ne mentionnent pas la durée des énoncés. Or bien que les énoncés de

MORA (1990) soient passablement plus courts que ceux de BYRD (1994), le résultat est pourtant le même : les hommes parleraient plus vite que les femmes. La durée des énoncés ne semble donc pas devoir être la cause des différences de débit. Ces quelques comparaisons suffisent à souligner la complexité de l'interprétation de résultats pourtant tout à fait objectifs, quantifiés.

Il ressort de ce survol de quelques études générales sur le lien entre la prosodie et les trois facteurs sociaux ici retenus, le sexe, l'âge et le statut socioprofessionnel que :

(i) un même phénomène intonatif (dans les études en cause, il s'agissait d'un contour terminal déclaratif montant) entraîne des interprétations variables selon les chercheurs;

(ii) un même phénomène intonatif (ici la modulation de l'énoncé) entraîne des résultats différents selon le type de corpus dans lequel il est examiné;

(iii) le traitement socioprosodique du phénomène intonatif apparaît plus délicat que le traitement du débit; en tout cas, dans les études examinées, les résultats à propos du débit paraissent plus convergents.

La prochaine section rappelle le type de données et la méthode de classification de ces données : une classification d'abord syntaxique, /discours direct/ et discours indirect, puis une classification discursive, selon la valeur narrative.

3.5.2 LA DISTRIBUTION DES ÉNONCÉS

Les sujets

On se rappelle que le corpus est issu de la production de 16 locuteurs représentant 8 hommes et 8 femmes, 8 locuteurs âgés de 15 à 29 ans et 8 locuteurs âgés de 51 à 73 ans ainsi que 8 locuteurs appartenant au groupe socioprofessionnel faible ou moyen et 8 locuteurs appartenant au groupe socioprofessionnel élevé (pour le rappel des détails, cf. section 3.1.2).

416

Il aurait été intéressant de faire une analyse multivariée des trois facteurs sociaux; par exemple, comparer les hommes âgés à statut socioprofessionnel élevé aux hommes âgés à statut socioprofessionnel faible et moyen. Toutefois, chacune de ces catégories n'est ici représentée que par 2 locuteurs et parfois même par un seul locuteur (1 femme jeune à statut socioprofessionnel faible ou moyen et 1 femme aînée à statut socioprofessionnel élevé). Il est alors apparu qu'il aurait été hasardeux de fonder des conclusions sur un nombre aussi limité de locuteurs. L'analyse porte donc sur l'effet de chacun des facteurs pris séparément.

Les énoncés

Le corpus se compose de 215 énoncés de discours rapporté tirés d'entrevues. La longueur des énoncés varie entre 4 et 31 syllabes, pour une médiane de 13 syllabes. Le regroupement des énoncés se fait en fonction de paramètres linguistiques, syntaxique et discursif, ainsi qu'en fonction de paramètres sociaux, le sexe, l'âge et le statut socioprofessionnel. Le *Tableau 3.13* en montre la distribution.

Tableau 3.13 Distribution des énoncés de discours rapporté selon le type de classification et selon les facteurs sociaux. F (femmes), H (hommes), J (jeunes), A (aînés), F ou M (faible ou moyen), É (élevé). R (*reproduction*), Act (*actualisation*).

	SEXE		ÂGE		STATUT	
	F	H	J	A	F ou M	É
	% (N)	% (N)	% (N)	% (N)	% (N)	% (N)
CLASSIFICATION SYNTAXIQUE						
/DD/ (n=163)	57% (93)	43% (70)	56% (91)	44% (72)	48% (79)	52% (84)
DI (n=52)	67% (35)	33% (17)	46% (24)	54% (28)	56% (29)	44% (23)
CLASSIFICATION DISCURSIVE						
/DD/ (n=163)						
R (n=41)	49% (20)	51% (21)	41% (17)	59% (24)	51% (21)	49% (20)
Act (n=68)	57% (39)	43% (29)	56% (38)	44% (30)	44% (30)	56% (38)
Autres (n=54)	48% (26)	52% (28)	72% (39)	28% (15)	50% (27)	50% (27)
DI (n=52)						
R (n=23)	70% (16)	30% (7)	57% (13)	43% (10)	43% (13)	57% (10)
Act (n=10)	80% (8)	20% (2)	20% (2)	80% (8)	70% (7)	30% (3)
Autres (n=19)	58% (11)	42% (8)	47% (9)	53% (10)	46% (9)	53% (10)

Comme on peut le constater dans le *Tableau 3.13*, la distribution des énoncés selon les facteurs sociaux, particulièrement en discours indirect, et encore davantage dans la classification discursive, occasionne le fractionnement des données. Le nombre restreint d'énoncés rendrait fragile la poursuite d'une étude statistique des contrastes entre le /discours direct/ et le discours indirect. De ce fait, chaque type de discours est analysé séparément.

118

Les éléments prosodiques considérés sont la configuration de F_0 à la transition[26] entre le discours citant et le discours cité, le débit en termes de vitesse d'articulation, c'est-à-dire sans les pauses silencieuses, ainsi que l'occurrence et la durée de la pause silencieuse à la transition.

3.5.3 LES HYPOTHÈSES

L'analyse prosodique de la production du discours rapporté ainsi que l'examen des études socioprosodiques précédentes amènent à poser les hypothèses suivantes.

1. Compte tenu que l'intonation du discours rapporté apparaît surtout liée à la syntaxe dans le cas du discours indirect et liée à la valeur discursive dans le cas du /discours direct/,
   l'hypothèse suivante est posée : la distinction intonative /discours direct/ et discours indirect serait peu corrélée à des facteurs sociaux.

2. Compte tenu que la pause silencieuse à la transition est hautement corrélée à la syntaxe (absente en discours indirect et plus souvent présente en discours direct lorsque le verbe *dire* est absent) et que, de plus, une étude générale sur la relation entre l'utilisation de la pause et le sexe (BYRD:1994) ne démontre aucune relation,
   l'hypothèse suivante est posée : la distinction pausale /discours direct/ et discours indirect serait peu corrélée à des facteurs sociaux.

---

[26] La relation entre les domaines du *groupe intonatif avant la transition* et du profil global de l'énoncé avec les facteurs sociaux ne s'est pas avérée significative; elle n'est donc pas présentée.

3. Compte tenu que (i) le débit s'est avéré plus ou moins rentable dans la distinction /discours direct/[27] et discours indirect (cf. toutefois *Figures 3.7* et *3.8* et *3.11*), que (ii) par ailleurs, des études (RYAN:1972; RAMIG:1983; MORA:1990; BYRD:1994; van RIE *et al.*:1995) mettent en relation le débit et des facteurs sociaux (sexe et âge),
l'hypothèse suivante est posée : (i) la distinction relative de débit entre le /discours direct/ et le discours indirect serait peu corrélée à des facteurs sociaux; (ii) en revanche, le débit du discours rapporté en général serait corrélé à des facteurs sociaux.

3.5.4 LES RÉSULTATS

Les résultats sont présentés selon les deux grands types de classification des énoncés: la classification syntaxique et la classification discursive.

À partir de la classification syntaxique, il ressort que :

1) **La configuration de F_0 à la transition**. En /discours direct/, la configuration de F_0 à la transition ne présente aucune corrélation avec les facteurs sociaux. En discours indirect, il y a corrélation avec le sexe ($p < 0,05$) et avec l'âge ($p < 0,01$) : la chute à la transition est plus marquée chez les hommes et chez les aînés.
2) **Le débit**. En /discours direct/, la vitesse d'articulation médiane de l'énoncé est plus élevée chez les hommes ($p < 0,01$) ainsi que dans le groupe socioprofessionnel élevé ($p < 0,05$). En discours indirect, cette vitesse d'articulation est aussi plus élevée chez les hommes ($p < 0,01$); chez les jeunes, la tendance est moins marquée ($p = 0,054$).

---

[27] Certes, pour le débit, il aurait été plus satisfaisant de distinguer les deux sous-catégories de discours direct (avec et sans verbe *dire*), mais les conséquences de la multiplication des subdivisions y a fait renoncer.

3) **La pause silencieuse**. Ni l'occurrence, ni la durée des pauses silencieuses à la transition ne sont liées à un facteur social.

À partir de la classification discursive, il ressort que :

1) **La configuration de F_0 à la transition**. En discours indirect, dans la classification discursive comme dans la classification syntaxique, la marque de la transition est plus forte chez les aînés ($p < 0,05$), particulièrement dans le type *reproduction*.
2) **Le débit**. En /discours direct/, dans la classification discursive comme dans la classification syntaxique, le débit médian est plus rapide chez les hommes ($p < 0,05$) ainsi que chez les locuteurs à statut socioprofessionnel élevé ($p < 0,05$) dans le type *actualisation*; dans le type *autres*, il est aussi plus rapide chez les hommes ($p < 0,05$). En discours indirect, les hommes ont également un débit plus rapide que les femmes dans le type *autres* ($p < 0,05$); dans le type d'emploi *reproduction*, les jeunes ($p < 0,05$) ainsi que les locuteurs du groupe favorisé ($p < 0,01$) ont un débit plus rapide.
3) **La pause silencieuse**. Dans la classification discursive, le trop petit nombre de cas dans chaque type d'emploi discursif rend impossible l'analyse de ce paramètre.

3.5.5 DISCUSSION

Corrélations observées

Le *Tableau 3.14* fait état des relations statistiquement significatives entre le sexe, l'âge et le statut socioprofessionnel d'une part et la configuration de F_0 à la transition, le débit et la pause silencieuse d'autre part, à partir d'une classification syntaxique puis discursive. Il reste à examiner dans quelle mesure ces corrélations relèvent de la distinction des types de discours rapporté.

Tableau 3.14 Corrélations entre facteurs sociaux et facteurs prosodiques dans la présente étude.

	F_0 À LA TRANSITION		DÉBIT (de l'énoncé)		
	Sexe	**Âge**	**Sexe**	**Âge**	**Statut**
/DD/					
C. syntaxique	∅	∅	H	∅	É
C. discursive	∅	∅	H**	∅	É**
			H*		
DI					
C. syntaxique	H	A	H	∅	∅
C. discursive	∅	A***	H*	J***	É***

H = homme; É = élevé; A = aînés; J = jeunes
***Reproduction;* ** *Actualisation;* * *Autres*

Il ressort de la corrélation entre les facteurs prosodiques et les facteurs sociaux retenus pour la caractérisation du discours rapporté les éléments suivants :

1) Le marquage intonatif du /discours direct/ n'est pas corrélé aux facteurs sociaux. En discours indirect, le marquage intonatif apparaît corrélé au facteur sexe : il est plus marqué chez les hommes (mais la comparaison porte sur 17 énoncés produits par des hommes par rapport à 35 produits par des femmes) et au facteur âge: il est plus important chez les locuteurs âgés, dans les deux types de classification. À partir de la classification syntaxique, la chute à la transition présente un écart médian de -5,75 DT (28 énoncés produits par les aînés) par rapport à -1,1 DT (24 énoncés produits par les jeunes). À partir de la classification discursive, dans le type d'emploi de la *reproduction*, la chute à la transition présente un écart médian de -5,7 (10 énoncés produits par les aînés) par rapport à -0,8 (13 énoncés produits par les jeunes). Est-ce que cette dernière corrélation entre l'intonation du discours indirect et l'âge est véritablement distinctive des deux types de discours rapporté ?

2) Tel qu'attendu, le débit est le paramètre prosodique le plus corrélé aux facteurs sociaux, mais la distinction relative de débit entre /discours direct/ et discours indirect apparaît peu corrélée aux facteurs sociaux. En effet, dans les deux classifications et dans les deux types de discours rapporté, les hommes ont régulièrement un débit plus rapide que les femmes; de même, les locuteurs appartenant à un groupe socioprofessionnel élevé ont parfois un débit plus rapide que ceux appartenant aux groupes socioprofessionnels faible et moyen. Il reste qu'en discours indirect, dans le type d'emploi discursif de la *reproduction*, les jeunes (13 énoncés) ont un débit plus rapide (6,48 syll./s par rapport à 5,68 syll./s) que les aînés (10 énoncés). Est-ce que cette dernière corrélation entre le débit du discours indirect et l'âge est véritablement distinctive des deux types de discours rapporté ?

3) Dans la classification syntaxique, la pause silencieuse n'est corrélée à aucun facteur social. Dans la classification discursive, en raison du nombre limité d'énoncés, l'étude n'a pu être menée.

Corrélations présentes dans les deux types de discours rapporté

Lorsque la corrélation entre une caractéristique prosodique et un facteur social est présente dans les deux types de discours rapporté et qu'en plus, elle a déjà été observée dans des études qui ne portaient pas sur le discours rapporté, cette corrélation n'est pas distinctive des deux types de discours rapporté. C'est le cas de la relation entre le débit et le sexe (plus rapide chez les hommes) ainsi qu'entre le débit et le statut socioprofessionnel (plus rapide chez les locuteurs du groupe socioprofessionnel élevé).

(i) Débit et sexe

Les résultats de la présente étude montrant une corrélation entre le débit et le sexe correspondent à ceux de MORA (1990) sur l'espagnol vénézuélien ainsi qu'à ceux de BYRD (1994) sur l'anglais américain, mais s'opposent à ceux de RYALLS *et al.* (1994) sur le français canadien. L'explication de cette corrélation semble rattachée au fait que les femmes ont un langage plus soigné que les hommes. COHEN (1989) a d'ailleurs

démontré que les phénomènes de réduction étaient plus fréquents chez les hommes que chez les femmes. On peut supposer, mais sans l'affirmer, que le fait d'utiliser une langue moins soignée, avec des formes réduites plus nombreuses, explique que les hommes aient un débit plus rapide que les femmes. Mais pourquoi les hommes ont-ils tendance à utiliser une langue moins soignée que les femmes ? À mon avis, il s'agit là d'une question sans réponse objective...

Quoi qu'il en soit, en discours rapporté, la corrélation entre le sexe et le débit ne permet pas de distinguer les types de discours rapporté; elle ne fait que refléter une tendance générale déjà observée dans d'autres études.

(ii) Débit et statut socioprofessionnel

En /discours direct/ comme en discours indirect, il apparaît que le débit est plus rapide chez les locuteurs appartenant au groupe socioprofessionnel élevé. On peut évoquer les mêmes explications que pour la corrélation avec le sexe et aboutir à la même question sans réponse. Cette corrélation n'est cependant pas signalée dans les études consultées. Par ailleurs, comme elle est commune aux deux types de discours rapporté, elle ne peut constituer un signe distinctif des catégories de discours rapporté.

Corrélations présentes dans un seul type de discours rapporté

Lorsque la corrélation entre une caractéristique prosodique et un facteur social n'est présente que dans un seul type de discours rapporté, il faut se demander si cette corrélation a déjà été observée dans d'autres études, pour d'autres types de structure. Si tel est le cas, la corrélation pourrait également être distinctive des deux types de discours rapporté. Rappelons que cette catégorie de corrélation (présente dans un seul type de discours rapporté) se produit en discours indirect, structure dont on a déjà signalé la ressemblance syntaxique et intonative avec l'énoncé neutre.

(i) Intonation et sexe

La corrélation entre l'intonation et le sexe montre ici que les hommes marquent plus la structure syntaxique du discours indirect que les femmes. Cette relation entre l'intonation et le sexe a déjà été observée par d'autres chercheurs dont BREND (1972), McCONNELL-GINET (1978) et RYALLS *et al.* (1994). Cependant, dans ces études faites à partir de corpus lus, l'aspect «expressif» relevé par BREND (1972) et McCONNELL-GINET (1978) chez les femmes ainsi que le marquage intonatif plus important d'une structure syntaxique (question par rapport à affirmation) par les femmes (RYALLS *et al.*:1994) se retrouvent plutôt chez les hommes dans les énoncés de discours rapporté. En revanche, dans une étude de MAURY et WRENN (1973:106), faite à partir d'un corpus spontané, ce sont également les hommes qui marquent davantage la différence entre la phrase interrogative et la phrase énonciative, comme dans la présente étude pour la structure syntaxique du discours indirect (une transition descendante plus forte). Il faut cependant ajouter que dans l'étude de MORA (1990), pourtant faite à partir d'un corpus spontané, aucune relation n'a été établie entre la modulation et le sexe.

Au-delà du fait que la modulation d'une structure puisse être reliée au sexe, et qu'un même type de corpus semble donner un même type de liaison (en lu, les femmes modulent plus; en spontané, ce sont les hommes), comment expliquer ces faits ? Des explications comme celles de McCONNELL-GINET (habitude du contact avec les enfants et de la situation de subordonnée) ne conviennent pas aux résultats obtenus avec des corpus spontanés. En outre, il semble hasardeux d'affirmer que l'«expressivité», prise dans sa globalité, n'est reliée qu'au facteur sexe. Le contexte est à tout le moins aussi important.

Ceci dit, il est difficile d'affirmer que la corrélation entre l'intonation du discours indirect et le sexe est distinctive des deux types de discours rapporté, parce que le nombre d'énoncés ici comparés est relativement restreint et très différent chez les hommes (17 énoncés) et chez les femmes (35 énoncés).

(ii) Intonation et âge

Que la classification des discours rapportés soit syntaxique ou discursive, il apparaît également que les aînés marquent davantage la transition du discours indirect. Par ailleurs, d'autres études ont montré une association entre l'intonation et les facteurs sexe et âge, comme c'est le cas dans la présente étude. Mais pourquoi les hommes, âgés, présenteraient-ils une pente plus forte que les autres types de locuteurs en discours indirect?

Selon MYSAK (1959) et HOLLIEN et SHIPP (1972), la fréquence usuelle augmente chez les hommes à partir d'environ 65 ans alors qu'elle diminue chez les femmes. Van RIE et van BEZOOIJEN (1995) attribuent le phénomène à des raisons physiologiques : il y aurait modification des cordes vocales. Ils ajoutent cependant qu'il est également possible que les personnes âgées se conforment à une idée stéréotypée de ce que doit être la voix d'une personne âgée. On peut ici se rappeler l'étude de SACHS (1975) avec des pré-adolescents; ceux-ci adaptent leur voix à leur sexe, avant que des modifications physiologiques ne se produisent réellement.

Dans le cas présent, la raison physiologique ne peut être soutenue, puisque les conséquences de l'augmentation du registre chez les hommes âgés par rapport à la diminution chez les femmes âgées ne se manifestent pas en /discours direct/. S'agit-il alors d'une distinction /discours direct/ et discours indirect liée aux facteurs sexe et âge ? C'est possible, mais il serait imprudent de l'affirmer.

(iii) Débit et âge

Dans le type d'emploi de la *reproduction*, le débit des locuteurs jeunes est plus rapide que celui des aînés. Cette relation entre le débit et le l'âge (plus rapide chez les jeunes) a déjà été observée, entre autres par RYAN (1972) et RAMIG (1983). Il peut être supposé que les organes articulatoires se meuvent moins rapidement au fur et à mesure que l'on avance en âge –encore faudrait-il tenir compte de l'état de santé des locuteurs observés. Si tel était le cas, le phénomène se

produirait aussi en /discours direct/. Faut-il alors croire à une distinction /discours direct/ et discours indirect corrélée à ces facteurs sociaux ?

Cette relation entre le débit et l'âge n'a été observée que pour un type d'emploi discursif représentant un petit nombre d'énoncés (13 énoncés produits par des jeunes et 10, par des aînés). Il apparaît téméraire de la considérer comme significative de la distinction /discours direct/ et discours indirect, ou d'une valeur narrative forte.

Dans l'ensemble, il est difficile d'affirmer qu'il y a corrélation entre les marques prosodiques du discours rapporté et les facteurs sociaux étudiés.

3.6 CONCLUSION

D'une part, le discours indirect est le type de discours rapporté qui présente les caractéristiques prosodiques les plus stables : montée dans la partie citante, chute dans la partie citée, jusqu'à la fin de l'énoncé, comme un énoncé neutre; pas de «marqueurs transitionnels» entre la partie citante et la partie citée. Il est à remarquer qu'il en est ainsi du point de vue syntaxique, le discours indirect contient toujours un verbe de parole et il est rare que la conjonction de subordination *que* soit absente (3 cas sur les 52 étudiés). L'organisation syntaxique du discours indirect est tout à fait comparable à celle de la complétive, et son organisation prosodique également. Du point de vue discursif, la valeur narrative n'influence pas véritablement les tendances fréquencielles du discours indirect.

D'autre part, les /discours directs/ ont des caractéristiques prosodiques plus fluctuantes, bien qu'elles soient généralement contrastives par rapport à celles du discours indirect : en effet, la partie citante est plus souvent descendante ou plate que montante, la

transition est plus souvent montante ou plate que descendante, le profil global de l'énoncé est plus souvent montant ou plat que descendant, les «marqueurs transitionnels» sont courants entre la partie citante et la partie citée. Cette variabilité prosodique se reflète à des degrés divers dans les structures syntaxique et discursive. La construction syntaxique du /discours direct/ est instable : tantôt un verbe *dire* introduit le discours cité, tantôt aucun verbe n'introduit le discours cité; tantôt un «marqueur transitionnel» sépare la partie citante de la partie citée, tantôt il n'y a pas de «marqueur transitionnel» entre la partie citante et la partie citée. Pourtant, la corrélation entre la fluctuation des caractéristiques prosodiques et cette fluctuation syntaxique apparaît relativement faible. En effet, même si l'absence de verbe *dire* semble occasionnellement comblée par la présence d'une pause silencieuse à la transition (28% en DDø par rapport à 11% en DD) ou par un maintien du débit entre le discours citant et le discours cité (par rapport à un changement de débit en DD et en DI), cette absence de verbe *dire* n'a pas d'effet sur la ligne mélodique de l'énoncé, pas plus que la présence de marqueurs d'attaque. En revanche, le type d'emploi discursif de la *reproduction* est corrélé à la transition montante (cf. *Figure 3.10;* $p < 0,05$) ainsi qu'à un débit moyen d'énoncé plus rapide (cf. *Figure 3.11;* $p < 0,01$). Les variations prosodiques observées en /discours direct/ semblent davantage reliées à des différences dans la valeur discursive de l'énoncé qu'à des différences dans la construction syntaxique de l'énoncé.

D'un point de vue socioprosodique, les corrélations entre facteurs prosodiques et facteurs sociaux sont bien réelles (observées dans plusieurs langues et dans plusieurs recherches), mais la nature fluctuante de la prosodie du spontané —probablement plus particulièrement de l'intonation— ainsi que la variété des méthodes utilisées exigent circonspection dans l'interprétation des résultats. Ceux-ci doivent être corroborés par plus d'une étude, les méthodologies doivent être comparables et les explications, basées sur des faits objectifs. En ce qui concerne le discours rapporté, globalement, il y a peu de corrélation entre les facteurs sociaux retenus et la prosodie distinctive des deux catégories de discours rapporté.

Contrairement aux tendances prosodiques nettes obtenues en corpus lu (LÉANDRI 1993), en corpus spontané, les pourcentages de récurrence des tendances prosodiques ne sont pas très élevés (6 fois sur 10 dans le meilleur des cas). La délimitation des groupes intonatifs ne correspond pas toujours à la séparation syntaxique discours citant/discours cité, pas plus qu'au type d'emploi discursif. Par surcroît, on peut même trouver des discours directs porteurs des caractéristiques prosodiques *typiques* du discours indirect.

Ceci dit, même si les rencontres prosodie-syntaxe-discours semblent plus fréquentes que ce que le seul hasard pourrait provoquer, il est difficile d'établir les liens forts entre ces niveaux linguistiques. La prosodie peut parfois s'allier à la structure syntaxique; ce sont les cas que l'on retrouve dans les tendances, particulièrement en discours indirect. Elle peut aussi quelquefois suppléer à la structure syntaxique, comme par exemple la présence relativement plus fréquente de la pause silencieuse à la transition des discours directs introduits sans verbe *dire*. Néanmoins pour l'ensemble des /discours directs/, la prosodie reflète davantage une valeur discursive. En effet, lorsque la valeur narrative est à son plus haut degré (type d'emploi de la *reproduction*), il y a une augmentation significative des occurrences de transitions montantes et le débit moyen de l'énoncé est significativement plus rapide. Toutefois, chaque niveau linguistique est à la fois trop fort pour ne répéter que ce que l'autre niveau a déjà permis de dire et trop fragile pour tout dire à lui seul. Le va-et-vient d'un niveau à l'autre apparaît constant, inévitable et ... pratiquement imprévisible.

Dans la littérature, on retrouve d'ailleurs, à maintes reprises, l'idée que la syntaxe et la prosodie ne s'associent que de façon bien relative. BOLINGER (1958) affirme que «encounters between intonation and grammar are casual, not causal». Pour CRUTTENDEN (1970), les croisements entre les deux sont simplement *derivative*, c'est-à-dire que la rencontre occasionnelle de la syntaxe et de la prosodie est davantage liée à une valeur sémantique plus générale des tons qu'à une structure syntaxique particulière. BARRY (1981) parle pour sa part de relations

coincidental, ce qui revient à la position de BOLINGER. COUPER-KUHLEN (1986) soutient que si le locuteur peut choisir de faire porter à l'intonation une valeur informative, grammaticale, «attitudinale», etc., c'est bien la preuve que l'intonation joue un rôle important dans le langage, mais que pour aucune de ces fonctions prises isolément ne peut être revendiquée une **version forte**.

Cette analyse de la production du discours rapporté a montré que la prosodie du discours rapporté est parfois liée à la forme syntaxique de l'énoncé, mais pas toujours, que la prosodie est aussi parfois liée à la valeur discursive de l'énoncé, mais pas toujours... La variabilité de ces associations présente une réalité, «non [d']un aveu d'impuissance» comme a déjà dit FRANÇOIS (1966).

Chapitre 4

ANALYSE DE LA PERCEPTION

La question délicate que se pose souvent le phonéticien au terme d'une analyse de la production : la machine «a saisi» les éléments produits qui viennent d'être décrits, mais l'oreille de l'interlocuteur, qu'a-t-elle «entendu» ? Et c'est bien sûr ce qui compte linguistiquement, on ne parle pas à des machines, mais à nos semblables ! L'objectif du chapitre 4 est de vérifier dans quelle mesure ces caractéristiques sont perçues par l'auditeur. Dans la présente étude, le seuil de perception des écarts intersyllabiques a été établi à 2 DT pour les intervalles mineurs (MERTENS 1993:23) et à 4 DT pour les intervalles majeurs (cf. section 2.2.1). Le seuil de perception du mouvement de F_0 plus étendu que celui entre deux syllabes a été fixé à 2 DT, puis à 3 DT (cf. section 3.2.1). Sur ces bases, des caractéristiques prosodiques dites typiques ont été attribuées au /discours direct/ et au discours indirect.

Dans la première partie, tous les énoncés du corpus (le corpus original) sont soumis à des auditeurs. Toute l'information linguistique disponible sur l'enregistrement est donnée, à savoir l'information prosodique et segmentale. L'objectif est de mettre en relation le pourcentage d'auditeurs qui «reconnaîtront» le type de discours rapporté, direct ou indirect, (désigné comme le pourcentage de réussites) et les caractéristiques prosodiques dégagées à partir de l'analyse de la production.

Dans une deuxième partie, une dizaine d'énoncés représentatifs de discours typiques et atypiques (un sous-corpus) sont présentés sous diverses formes : avec l'information prosodique seulement, avec l'information segmentale seulement puis avec l'information prosodique modifiée. Ainsi, dans un premier temps, un seul type d'information est donné, soit une mélodie sans mots ou des mots sans mélodie. Dans un deuxième temps, les paramètres prosodiques sont modifiés, la fréquence fondamentale seule, la pause silencieuse, puis les deux simultanément. L'objectif est alors double : (i) vérifier, sur des énoncés choisis en fonction de leur typie et de leur atypie prosodique, l'importance de l'information prosodique dans la perception du discours rapporté; (ii) vérifier l'importance de deux paramètres prosodiques pris isolément, puis conjointement : la fréquence fondamentale et la pause silencieuse.

La perception du discours rapporté est analysée surtout sur une base grammaticale, /discours direct/ et discours indirect, puis parallèlement sur une base discursive, la valeur narrative. La perception de la valeur discursive est aussi examinée sur une base «attitudinale», à travers l'attitude d'*accord* ou de *désaccord* du rapporteur avec les paroles qu'il rapporte[1]. Par exemple, dans un énoncé comme :

(1) Mon mari il dit : « Ah ben c'est correct.» (5612-108),

est-ce que l'intonation permet de percevoir si la femme est d'*accord*, ou non, avec son mari ?

---

[1] L'objectif n'est pas de vérifier la perception des diverses valeurs discursives associées à diverses attitudes éventuellement liées au discours rapporté (tel AUTHIER:1978), mais simplement de vérifier dans quelle mesure la mélodie peut effectivement entraîner la perception d'une attitude spécifique, ici l'*accord* ou le *désaccord*, et dans quelle mesure les éléments prosodiques étudiés sont des éléments actifs dans la perception de cette attitude.

PARTIE A : LE CORPUS ORIGINAL[2]

4.1 LA MÉTHODOLOGIE

L'analyse du corpus original met en relation le pourcentage de réussites dans la perception de chacun des énoncés avec les paramètres prosodiques étudiés lors de l'analyse de la production :

1) la configuration de F_0 pour le *groupe intonatif avant la transition*, la transition, le *groupe intonatif après la transition* et le profil global de l'énoncé;
2) la présence des «marqueurs transitionnels»;
3) la différence de débit (vitesse de locution) entre le discours citant et le discours cité.

4.1.1 LE TYPE DE TEST

La nécessité de soumettre l'entier du corpus (215 énoncés) à un nombre important de sujets (autour d'une quarantaine), toujours les mêmes, requérait un type de test simple qui ne devait pas dépasser 4 séances d'environ 10 minutes chacune. Le type de test à réponse binaire, qui consiste ici à distinguer si l'énoncé appartient au discours direct ou au discours indirect (choix forcé), s'est finalement imposé. Un test à réponse indirecte n'a pas été jugé nécessaire; d'ailleurs, la notion de discours rapporté semble avoir été bien intégrée par les participants (cf. ci-après, section 4.1.3, Alpha de Cronbach). On trouvera à l'*Annexe A* un exemplaire de ce qui a été présenté aux sujets: l'explication, les consignes et un exemple de feuille-réponse.

---

[2] Les bandes sonores du corpus original sont disponibles sur demande.

4.1.2 LE PROTOCOLE

Afin de ne pas fatiguer les sujets, l'audition des 215 énoncés est répartie en 4 séances. La première dure 15 minutes et les suivantes, 10 minutes chacune. Les deux premières séances se déroulent la même journée : l'une au début du cours[3], l'autre, à la fin du cours, soit avec un intervalle d'un peu plus de deux heures. Les deux dernières séances se déroulent la semaine suivante, de la même façon : une séance en début de cours et une séance en fin de cours. Chaque séance permet de présenter 55 énoncés[4] tirés d'entrevues de 4 à 5 locuteurs différents (selon les séances). Étant donné la qualité tout juste moyenne des enregistrements ainsi que la fragilité sémantique de paroles sorties de leur contexte, il a été décidé de présenter à la suite tous les énoncés d'un locuteur afin de maximiser l'effet d'adaptation à la qualité d'une voix. L'ordre de présentation des locuteurs a suivi l'ordre numérique des locuteurs (locuteur 15, locuteur 30, locuteur 59, etc.), l'ordre de présentation des énoncés de chaque locuteur a été fixé aléatoirement. Chaque énoncé est répété 2 fois d'affilée; il y a une seconde d'intervalle entre les répétitions, 3 secondes entre les énoncés et 6 secondes entre les groupes de 5 énoncés. Les sujets sont prévenus qu'il entendront des locuteurs différents, qu'il s'agit d'énoncés extraits d'entrevues et que le débit de certains énoncés est rapide.

4.1.3 LES SUJETS

Les sujets sont des étudiants de premier cycle universitaire, de langue maternelle franco-québécoise. Du fait que l'expérience se déroule en 4 séances et dans deux groupes parallèles (Groupes 1 et 2), le nombre de sujets varie légèrement d'une séance à l'autre.

---

[3] L'expérience se déroule dans le cadre d'un cours universitaire.

[4] Pour contrer l'effet de «surprise» lié aux 5 premiers énoncés de la première série, ceux-ci ont été présentés une nouvelle fois à la fin de la quatrième série.

Afin de vérifier le degré de fiabilité des sujets répondants, l'alpha de Cronbach (CRONBACH:1951) a été appliqué. Pour calculer l'alpha de Cronbach, il ne doit pas y avoir de valeurs manquantes. Les réponses des sujets qui ont été absents à une séance ou plus sont donc exclues; c'est le cas de 15 des 47 sujets. De plus, seuls les énoncés pour lesquels aucune réponse n'est *sans réponse* ou *annulée* ont été retenus; c'est le cas pour 157 des 215 énoncés. L'alpha de Cronbach pour les 32 sujets et les 157 énoncés est de 0,856, c'est-à-dire que, dans ces conditions, le degré de fiabilité est de 86%. La fiabilité des taux de perception issus de l'Expérience 1 est donc considérée comme satisfaisante.

4.2 LES HYPOTHÈSES GÉNÉRALES

D'une part, on peut faire l'hypothèse que le sujet repère le type de discours rapporté à partir de la structure syntaxique. D'autre part, dès qu'il y a de l'oral, les auditeurs sont très sensibles à la prosodie (MEHTA et CUTLER:1988). Dans une certaine mesure, il est probable que le sujet puisse repérer le type de discours rapporté à partir des caractéristiques prosodiques. Il s'agit ici de vérifier cette dernière hypothèse. Les caractéristiques prosodiques supposées favorables à la reconnaissance d'un type de discours rapporté découlent de l'analyse de la production. Les hypothèses générales sont les suivantes :

1. Compte tenu que les caractéristiques prosodiques du discours indirect sont les plus stables,
   l'hypothèse suivante est posée : le discours indirect serait mieux perçu que le discours direct.

2. Compte tenu que c'est le domaine de la transition qui représente le plus significativement le type de discours (cf. *Figure 3.3*),
   l'hypothèse suivante est posée : la transition serait un domaine majeur dans la perception des types de discours rapporté.

3. Compte tenu que la différence de débit entre le discours citant et le discours cité est semblable en discours indirect et en discours direct avec verbe *dire* (cf. *Figure 3.7*),
l'hypothèse suivante est posée : la différence de débit aurait peu d'effet sur le pourcentage de réussites dans la perception des types de discours rapporté.

4. a) Compte tenu que les «marqueurs transitionnels», particulièrement les marqueurs d'attaque et les pauses silencieuses, sont relativement fréquents en /discours direct/ et que de plus, dans le présent corpus, ils lui sont exclusifs (cf. *Tableau 3.4*),
l'hypothèse suivante est posée : ces «marqueurs transitionnels» auront un effet sur le pourcentage de réussites dans la perception des types de discours rapporté.

b) Compte tenu de cette fréquence d'occurrence des «marqueurs transitionnels» dans la perception du /discours direct/, mais de leur peu d'influence sur la ligne mélodique (cf. section 3.2.2),
l'hypothèse suivante est posée : en présence de «marqueurs transitionnels», la valeur d'écart de la transition aurait peu d'effet sur le pourcentage de réussites dans la perception des types de discours rapporté.

5. Compte tenu que certaines configurations tonales sont significativement cooccurrentes (cf. *Tableau 3.3*),
l'hypothèse suivante est posée : l'interaction entre ces configurations tonales aurait un effet sur le pourcentage de réussites dans la perception des types de discours rapporté.

6. Compte tenu que les caractéristiques prosodiques du /discours direct/ sont généralement plus corrélées à la valeur discursive de reproduction, de *(+) narrative* à *(-) narrative*, en /discours

direct/ qu'en discours indirect (cf. section 3.3.4),
l'hypothèse suivante est posée : les énoncés en /discours direct/ à forte valeur de reproduction, *(+) narrative*, seraient mieux perçus que ceux à faible valeur de reproduction, *(-) narrative*, alors que la perception des énoncés en discours indirect ne serait pas fonction de la valeur narrative des énoncés.

4.3 LES RÉSULTATS

EXPÉRIENCE 1 : INFORMATION PROSODIQUE ET SEGMENTALE

4.3.1 CLASSIFICATION SYNTAXIQUE

Il apparaît d'abord, par le simple calcul de la moyenne des pourcentages de réussite, que le discours indirect est globalement mieux perçu (73%) que le /discours direct/ (66,5%). Un premier modèle de régressions linéaires multiples, incluant les types de discours comme variables indépendantes, donne un R^2 de 9% dans les deux sous-catégories de discours direct (avec ou sans verbe *dire*) alors que le R^2 est de 11% en discours indirect. La différence entre le pourcentage de réussites en /discours direct/ et celui en discours indirect est significative ($p < 0,001$); (hypothèse 1).

Un second modèle capable d'expliquer le pourcentage de sujets percevant du /discours direct/[5] a ensuite été construit à l'aide de huit variables indépendantes, excluant les types de discours. Cinq sont des

[5] La modélisation a été effectuée pour le pourcentage de sujets qui ont perçu du discours direct (même en discours indirect). Ainsi, si un énoncé en discours indirect a été perçu par 72% des sujets comme du discours indirect, il reste 28% des sujets qui l'ont perçu comme du discours direct, et c'est ce 28% qui est entré dans le modèle. Les énoncés *sans réponse* et les *réponses annulées* sont exclus du dénominateur au même titre que les candidats absents.

variables continues : les mesures d'écarts pour le *groupe intonatif avant la transition*, la transition, le *groupe intonatif après la transition* et le profil global de l'énoncé (pour le rappel de la façon de mesurer, cf. *Figure 3.1*), ainsi que les différences de débit entre le discours citant et le discours cité (*Figure 3.7*). Trois sont des variables dichotomiques: la présence ou l'absence d'un marqueur d'attaque, la présence ou l'absence d'une pause silencieuse et la présence ou l'absence d'une pause d'hésitation.

Une première sélection des variables est effectuée à l'aide de régressions linéaires du pourcentage de sujets percevant du /discours direct/ sur chacune des huit variables explicatives séparément et sur chacune des interactions doubles formées à partir de ces huit variables. Seules les variables et interactions significatives au niveau de signification de 10% sont conservées. Elles sont incluses toutes ensemble dans un modèle de régression multiple qui est ensuite simplifié par l'élimination progressive des variables et interactions qui ne sont pas significatives dans la régression multiple. Le modèle final ne contient donc que les variables et interactions qui expliquent de façon significative le pourcentage de sujets percevant du /discours direct/. Cette approche est utilisée pour modéliser l'ensemble des 215 énoncés.

Pour les 215 énoncés, les variables significatives sont la transition, le marqueur d'attaque et la pause silencieuse (dans les 3 cas, $p < 0{,}001$); (hypothèses 2, 3 et 4). Par ailleurs, contrairement à ce qui était attendu, il n'y a pas d'interaction entre les différentes variables continues que sont les configurations tonales (hypothèse 5). On peut penser que le nombre d'énoncés (ce nombre varie de 5 à 15 énoncés) dans chacune des cooccurrences n'est pas suffisant pour que l'interaction ressorte comme significative dans le modèle.

Il y a cependant une interaction presque significative entre la transition et la présence d'une pause silencieuse (T*Ps; $p = 0{,}0627$). La *Figure 4.1* illustre cette interaction entre la transition et la pause silencieuse.

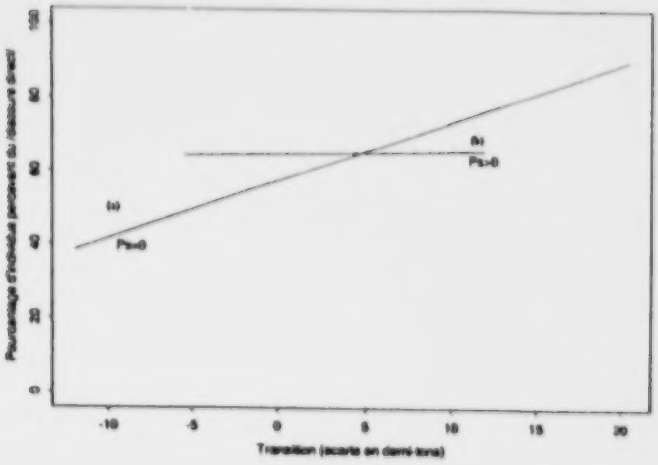


Figure 4.1 Corpus original (n = 215). Effet de la transition sur le pourcentage de sujets percevant du /discours direct/, selon l'absence (droite (a); Ps = 0) ou la présence d'une pause silencieuse (droite (b); Ps > 0).

Il faut d'abord préciser que la droite tracée pour illustrer l'effet de la variable sur un sous-groupe d'énoncés (avec ou sans pause silencieuse) est limitée aux valeurs observées de la variable dans ce sous-groupe. Par exemple, dans la *Figure 4.1*, la droite (b) illustre des transitions variant entre -5,56 DT et 11,94 DT. Chaque droite est accompagnée de sa pente *m* (la pente indique l'augmentation (pente positive) ou la diminution (pente négative) du pourcentage de sujets percevant du /discours direct/ pour chaque demi-ton) et de sa valeur *p*: droite (a) : m = 1,6 $p < 0{,}001$; droite (b) : m = 0,1 n.s.

La droite (a) de la *Figure 4.1* illustre que, lorsqu'il n'y a pas de pause silencieuse (c'est le cas pour 186 des 215 énoncés), chaque fois que l'écart augmente de 1 DT à la transition, le pourcentage de sujets

percevant du /discours direct/ augmente de 1,6%. Par exemple, le modèle permet de prédire qu'un énoncé sans pause silencieuse à la transition, dans lequel cette transition est égale à +2 DT (le seuil de perception fixé), sera perçu comme du discours direct par environ 60% des sujets.

La droite (b) illustre que, lorsqu'il y a une pause silencieuse à la transition, le pourcentage de perception en /discours direct/ est à peu près stable (≈ 65%). Ce pourcentage de perception est indépendant de l'importance de l'écart à la transition. La présence de la pause silencieuse devient le paramètre déterminant dans la reconnaissance du type de discours.

La *Figure 4.2* illustre une interaction significative entre le marqueur d'attaque et la pause silencieuse (MA*Ps; $p < 0,001$) à l'aide d'un histogramme, parce que les variables présentées sont dichotomiques.

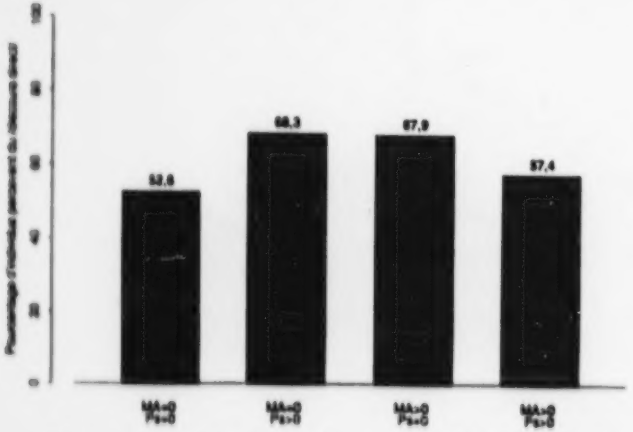


Figure 4.2 Corpus original (n =215). Pourcentage de sujets percevant du /discours direct/ selon l'absence (MA = 0) ou la présence (MA > 0) d'un marqueur d'attaque et d'une pause silencieuse (Ps = 0 ou Ps > 0)

La *Figure 4.2* indique que si on compare le pourcentage de perception lorsqu'il y a absence des deux variables (52,6%) au pourcentage de perception lorsque l'une des 2 variables est présente, la différence dans le pourcentage de perception est significative ($p < 0,001$), qu'il s'agisse de la pause silencieuse (Ps > 0; 68,3%) ou du marqueur d'attaque (MA> 0; 67,9%).

En revanche, si on compare le pourcentage de perception lorsqu'il y a présence des deux variables (57,4%) au pourcentage de perception lorsque l'une des 2 variables est présente, la différence dans le pourcentage de perception n'est pas significative, qu'il s'agisse de la pause silencieuse (68,3%) ou du marqueur d'attaque (67,9%).

Ainsi, le pourcentage de perception en /discours direct/ est non seulement augmenté par la présence d'un marqueur d'attaque ou d'une pause silencieuse, mais les deux sont en interaction de la façon suivante: le pourcentage de perception est favorisé lorsque seulement l'une des deux variables est présente. Le R^2 obtenu avec ce modèle est de 27,6%, ce qui est satisfaisant, compte tenu qu'il s'agit d'expliquer un pourcentage de perception.

4.3.2 CLASSIFICATION DISCURSIVE

Il est intéressant de signaler ce qui se passe du point de vue perceptuel lorsque les énoncés de discours rapporté sont classifiés sur une base discursive, c'est-à-dire selon leur valeur narrative : de la *reproduction, (+) narrative*, à l'*assertion, (-) narrative*, en passant par la *pseudo-reproduction*, l'*actualisation* et l'*invention* (valeurs narratives intermédiaires). Le *Tableau 4.1* montre les résultats

Tableau 4.1 Corpus original. Classification discursive. Pourcentage de réussites et types d'emploi discursif.

Types d'emploi discursif	DISCOURS DIRECT (n = 163)	% perc.	DISCOURS INDIRECT (n = 52)	% perc.
Reproduction	(n = 41)	75%	(n = 23)	75%
Pseudo-reproduction	(n = 14)	73%	(n = 1)	-
Actualisation	(n = 68)	63%	(n = 10)	82%
Invention	(n = 33)	65%	(n = 1)	-
Assertion	(n = 7)	58%	(n = 17)	69%

La démonstration faite dans l'analyse de la production, à savoir que les caractéristiques prosodiques sont, dans une certaine mesure, fonction de la valeur narrative en /discours direct/, mais pas en discours indirect, se répète du point de vue perceptuel (hypothèse 6). Le pourcentage de réussites dans la perception d'un énoncé est aussi fonction de la valeur narrative en /discours direct/, ce qui n'a rien de surprenant puisque les caractéristiques prosodiques sont plus marquées lorsque la valeur narrative est forte et vice-versa. Tel n'est pas le cas en discours indirect : les caractéristiques syntaxiques sont plus stables qu'en discours direct, les caractéristiques prosodiques aussi, et la valeur discursive ne modifie pas le pourcentage de réussites dans la perception d'un énoncé de forme indirecte.

4.4 CONCLUSION

La plupart des hypothèses générales posées sont confirmées par l'Expérience 1.

1. Le pourcentage de perception en discours indirect est supérieur à celui en /discours direct/.

Le modèle construit à partir des cinq variables continues et des trois variables dichotomiques montre que :

2. Tous types de discours confondus, la transition est la seule variable continue qui soit explicative du pourcentage d'individus percevant du /discours direct/ ($p < 0,001$).

3. La différence de débit entre le discours citant et le discours cité n'est pas une variable explicative.

4. a) Les variables dichotomiques que sont le marqueur d'attaque et la pause silencieuse sont explicatives du pourcentage de perception ($p < 0,001$ dans les deux cas).

   b) Tel qu'attendu, la présence d'une pause silencieuse suffit à elle seule à prédire un pourcentage de perception en /discours direct/, indépendamment de la valeur d'écart de la transition (T*Ps, *Figure 4.1*). La pause silencieuse constitue donc un paramètre prosodique important dans la perception. Par ailleurs, lorsqu'il n'y a pas de pause silencieuse, le pourcentage de perception est fonction de la valeur d'écart à la transition. Il apparaît que le seuil de perception de + 2 DT est largement suffisant puisque le pourcentage de perception atteint les 60% un peu avant que ce seuil ne soit atteint. D'un autre côté, le pourcentage de perception en /discours direct/ de l'énoncé sans pause silencieuse devient comparable à celui avec pause silencieuse lorsque l'écart à la transition atteint presque les 5 DT. Il y a une interaction non prévue entre le marqueur d'attaque et la pause silencieuse. Cette interaction (MA*Ps, *Figure 4.2*) montre que la présence de l'un ou l'autre augmente de façon significative le pourcentage de perception en /discours direct/ si on le compare au pourcentage obtenu lorsque ces deux variables sont absentes de l'énoncé.

5. Contrairement à l'hypothèse 5, il apparaît que les cooccurrences significatives des configurations tonales observées dans l'analyse de la production n'ont pas d'effet sur le pourcentage de perception en /discours direct/. Il est probable que le nombre restreint d'énoncés

présentant les cooccurrences significatives neutralise leur effet dans le modèle.

6. Tel qu'attendu, du point de vue discursif, le pourcentage de perception en /discours direct/ est plus élevé dans le type d'emploi discursif à valeur narrative forte que dans celui à valeur narrative faible alors que le pourcentage de perception en discours indirect n'est pas relié à cette valeur narrative.

Il ressort de cette Expérience 1 que la distinction perceptuelle entre /discours direct/ et discours indirect est non seulement liée à la structure syntaxique, mais aussi aux différentes caractéristiques discursives et prosodiques.

PARTIE B : UN SOUS-CORPUS[6]

4.5 LE SOUS-CORPUS

L'étude du sous-corpus porte principalement sur la fonction syntaxique et, secondairement, sur la fonction «attitudinale» de la prosodie à travers la structure syntaxico-discursive du discours rapporté.

Le sous-corpus a fait l'objet de 9 expériences: 7 concernent la différence syntaxique et 2, la différence d'attitude. Parmi les 7 qui portent sur la différence syntaxique, 4 utilisent l'information prosodique et/ou segmentale originale alors que dans les 3 autres, l'information prosodique est manipulée. Les 2 expériences qui portent sur la différence d'attitude sont faites à partir de l'information originale: prosodique et segmentale dans le premier cas, puis uniquement à partir de l'information prosodique dans le second.

---

[6] Les bandes sonores du sous-corpus sont disponibles sur demande.

4.5.1 LES ÉNONCÉS

Parmi les 215 énoncés analysés, 16 ont été retenus pour les tests de perception. Les paramètres de description ont été les suivants:

1. Représentativité des catégories étudiées de discours rapporté:
   - 12 /discours directs/[7] (12/16 soit 76% par rapport à 163/215 soit 75%),
   - 4 discours indirects (DI).
2. Représentativité de caractéristiques prosodiques dites typiques et atypiques :
   - 10 discours rapportés dits typiques (6 discours directs et 4 discours indirects),
   - 6 discours rapportés dits atypiques (6 discours directs).
3. Longueur des énoncés (entre 6 et 13 syllabes)[8].
4. Timbre et registre comparables. Les 3 critères précédents ont nécessité la sélection de 4 locuteurs du corpus (les locuteurs 15, 59, 79 et 108). Il s'agit de femmes âgées de plus de 50 ans et dont le registre moyen se situe entre 170 et 180 Hz.

4.5.2 LA DISTINCTION ÉNONCÉS DITS PROSODIQUEMENT TYPIQUES/ÉNONCÉS DITS ATYPIQUES

Il a été montré dans l'analyse de la production que les trois domaines prosodiques suivants, *groupe intonatif avant la transition*,

---

[7] Les discours directs ne sont pas ici réunis à partir de la présence ou de l'absence d'un verbe *dire*, mais plutôt à partir de la présence ou de l'absence de caractéristiques prosodiques typiques du /discours direct/.

[8] Comme la plupart des expériences portent sur les seules mélodies, il apparaissait nécessaire que celles-ci ne soient pas trop longues afin qu'elles ne soient pas hors de la capacité de décodage des sujets.

transition, et profil global de l'énoncé[9] sont significatifs (cf. *Tableau 3.3*). Les énoncés sont considérés prosodiquement typiques s'ils portent une transition conforme à la tendance : montante en /discours direct/ et descendante en discours indirect. Si la transition est plate, la présence d'une pause silencieuse est requise pour que l'énoncé soit associé à un discours direct typique. Les autres énoncés sont considérés prosodiquement atypiques.

Comme les énoncés de ce sous-corpus n'ont été choisis qu'en fonction de leurs caractéristiques prosodiques, il est intéressant de noter au passage les types discursifs représentés dans chaque catégorie. Les /discours directs/ dits prosodiquement typiques appartiennent au type discursif de la *reproduction* (6 sur 6). Les discours indirects dits prosodiquement typiques appartiennent respectivement au type *pseudo-reproduction, assertion, reproduction* et *actualisation*. Parmi les énoncés en /discours direct/ dits prosodiquement atypiques, 2 des 6 énoncés sont de type *reproduction*, 3 de type *actualisation* et 1 de type *assertion*. Les observations effectuées à propos de la répartition des types discursifs se reflètent donc dans ce sous-corpus : le type d'emploi discursif de la *reproduction* apparaît comme régulièrement associé à des caractéristiques prosodiques tendancielles. Il n'en est cependant pas tenu compte dans cette partie de l'analyse de la perception.

4.6 LA MÉTHODOLOGIE

L'objectif premier des tests de perception est de vérifier la part de l'information prosodique dans la reconnaissance de la structure syntaxique du discours rapporté alors que l'objectif secondaire est d'observer le rôle de l'intonation dans la reconnaissance de l'attitude (*accord* ou *désaccord*) du rapporteur par rapport aux paroles citées.

---

[9] La caractérisation du /discours direct/ par le profil montant est cependant fragile (cf. *Tableau 3.3*).

4.6.1 LES TYPES DE TESTS

Les types de tests doivent être variés, autant par l'information soumise que par la tâche demandée. Ainsi, on retrouve dans les expériences qui suivent trois types d'information : prosodique, segmentale ainsi que prosodique et segmentale à la fois et deux types de tâche : l'une consiste à associer un stimulus à la notion discours direct ou discours indirect (test-notion), l'autre consiste plutôt à associer un stimulus à la transcription en discours direct ou en discours indirect (test-transcription). Cette diversification de la tâche sert à mesurer le risque d'une maîtrise mal assurée de la notion proposée, c'est-à-dire à vérifier si la différence de performance est notable d'une tâche à l'autre.

Dans l'ensemble, le choix du type de tâche a été fonction du type d'information. Le matériau sans information segmentale est le plus souvent soumis à travers un type de test-transcription (cf. *Annexe B*), a priori plus simple, alors que le matériau porteur d'information segmentale utilise le test-notion (cf. *Annexe A*). Le caractère naturel, instinctif, du suprasegmental —acquis dans les premiers mois de vie (KONOPCZINSKI:1986 et 1990)— par rapport au caractère plus conscient du segmental est à la base de ce choix. Par ailleurs, à titre comparatif, deux expériences (3 et 10) avec information prosodique sont tout de même réalisées avec un type de test-notion.

4.6.2 LA TECHNIQUE DE «DÉLEXICALISATION»[10]

La première technique envisagée afin d'effacer le segmental a bien sûr été le filtrage des énoncés. Des essais faits avec les systèmes WAVE et LuPP[11] n'ont pas donné de résultats satisfaisants : ou l'intensité diminuait considérablement à la synthèse ou des bruits

---

[10] Les expériences (sauf l'Expérience 4) ont été conçues et réalisées à l'Institut de Linguistique et de Phonétique de Lund (Suède), avec la collaboration des professeurs Paul Touati et David House ainsi que de l'ingénieur Marcus Filippson.

[11] Lund Prosodic Parser

importants apparaissaient. Les signaux produits étaient difficilement acceptables pour des tests de perception

Il a alors été décidé de faire des essais avec la technique de «délexicalisation». De façon simple, la «délexicalisation» consiste à reproduire artificiellement un signal à partir de la fréquence du fondamental et de l'intensité. On voit ici l'intérêt d'appliquer cette technique à un corpus bruité : en plus d'effacer toute trace segmentale, les bruits disparaissent et l'intensité du signal est préservée, comme l'illustre la *Figure 4.3*.

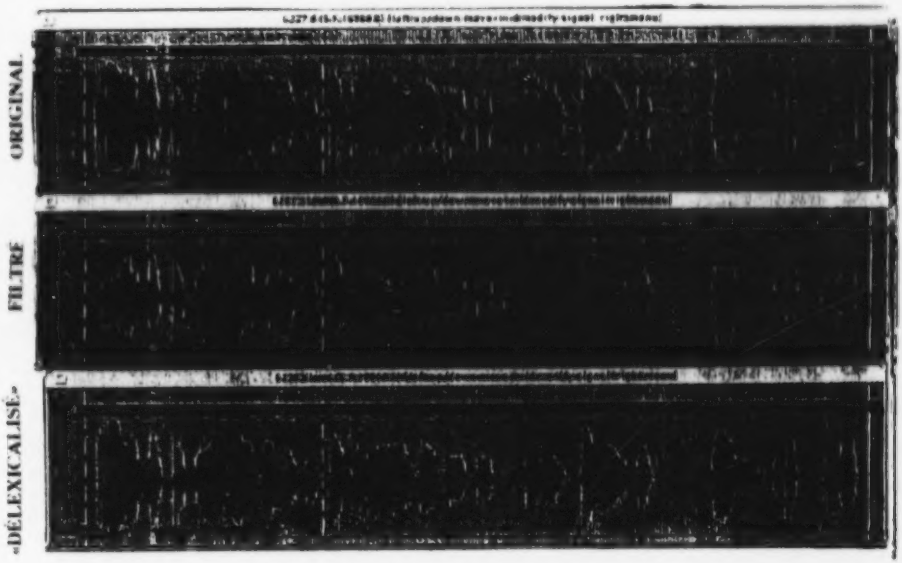


Figure 4.3 Signal original comparé à un signal filtré, puis «délexicalisé»

Le *hic* réside dans l'artificialité du signal ainsi produit. Plusieurs sortes d'ondes ont été testées, celles en dents de scie sont apparues à l'audition comme étant les moins artificielles. Les 16 signaux ont ainsi été «délexicalisés». Tous les signaux obtenus ont été jugés acceptables pour des tests de perception.

Une expérience (Expérience 4) a été réalisée avec le système de synthèse ASL[12]. Il s'agissait dans ce cas de monotoniser les énoncés, c'est-à-dire de supprimer l'information prosodique. La prosodie des énoncés a été ramenée à 180 Hz, fréquence de base des 3 locutrices du sous-corpus. La possibilité de supprimer les variations intersyllabiques, tout en gardant une certaine déclinaison de l'énoncé, aurait permis d'obtenir un énoncé moins «artificiel». Cependant, le fait qu'un profil global descendant soit tendanciel du discours indirect par rapport au profil montant du /discours direct/ a fait renoncer à ce type d'altération. L'énoncé qui résulte de la manipulation est donc systématiquement monotone. Par ailleurs, le système de synthèse utilisé produit une certaine forme d'écho à la limite de l'acceptable. Pour cette raison, un type de test à coefficient de difficulté moins élevé (répétitions consécutives) a été retenu pour l'Expérience 4.

4.6.3 LE PROTOCOLE

Afin de contrôler l'uniformité des explications et des consignes, celles-ci sont écrites au début de chaque test et lues aux sujets (cf. l'exemple de l'Expérience 1 en *Annexe A*). Les **répétitions consécutives**, au nombre de trois, se retrouvent dans les tests où il y a association de ce qui est entendu à la transcription d'un même énoncé sous les deux formes différentes du discours rapporté, le /discours direct/ et le discours indirect (Expériences 2, 6, 7, 8 et 9). Elles facilitent la tâche de décodage. Dans les tests où il y a association de ce qui est entendu avec une notion (Expériences 3, 5 et 10), il n'y a pas de répétitions consécutives, mais plutôt des **répétitions non consécutives**, au nombre de cinq, sauf dans l'Expérience 3 où deux répétitions ont été jugées suffisantes : il s'agissait

---

[12] L'Expérience 4 a été réalisée avec la collaboration du professeur Claude Paradis, au laboratoire de phonétique de l'Université Laval où le système disponible est ASL de Kay Elemetrics Corp.

d'énoncés avec l'entier de l'information, comme dans l'Expérience 1. Le but est de vérifier le degré de cohérence dans le maintien des choix[13]; ce procédé permet de mesurer la constance de l'ensemble des répondants pour chacun des stimuli. L'Expérience 4 fait exception. Les limites de la technique de synthèse utilisée ont requis des répétitions consécutives, même s'il s'agissait d'un type de test-transcription.

Quant à la longueur des intervalles, elle est fonction du type de répétition[14]. Lorsqu'il y a répétition consécutive, l'intervalle entre les répétitions est de 1 s et celui entre les stimuli de 8 s; lorsque les répétitions sont non consécutives, l'intervalle est de 3 s entre les stimuli et de 6 s entre les blocs de stimuli (4 ou 5 selon le nombre total de stimuli).

4.6.4 LES SUJETS

Règle générale, chaque groupe de sujets participe à 2 ou à 3 expériences de façon à ne pas dépasser une séance de 20 minutes. Toutes les expériences ont été soumises à des sujets naïfs. Ces sujets, entre 30 et 65, sont des étudiants du premier cycle universitaire, de langue maternelle franco-québécoise. Afin de vérifier si des connaissances en phonétique améliorent le pourcentage de réussites, trois (3) des expériences (Expériences 2, 7 et 8) ont également été soumises à des sujets experts, le Groupe 1ex. Les treize (13) sujets experts sont des étudiants de deuxième cycle en phonétique ainsi que des professeurs de phonétique.

---

13 La réponse d'un sujet est considérée cohérente lorsque la même réponse revient au moins 4 fois pour les 5 répétitions (ou 3 fois lorsque le sujet n'a répondu qu'à 4 stimuli). Le degré de cohérence est obtenu en divisant le nombre de sujets ayant des réponses cohérentes par le nombre de sujets ayant participé au test.

14 La longueur des intervalles a été fixée selon les conseils du professeur David House de l'Institut de Linguistique et de Phonétique de Lund.

4.7 LES RÉSULTATS

4.7.1 PERCEPTION DE LA DIFFÉRENCE SYNTAXIQUE

La fonction de la prosodie dans la distinction discours direct/discours indirect est ici appelée différence syntaxique. Cette fonction est étudiée à travers la perception d'une information prosodique et/ou segmentale originale (Expériences 2, 3, 4 et 5) ainsi que d'une information prosodique manipulée (Expériences 6, 7 et 8). Cinq des sept expériences utilisent la seule information prosodique, une expérience utilise la seule information segmentale et une autre, l'information prosodique et segmentale. Pour chaque grande catégorie d'information, non manipulée et manipulée, des hypothèses sont posées.

Le nombre restreint d'énoncés (en moyenne, une dizaine par expérience) rend l'analyse statistique difficile. Elle est utilisée à titre indicatif seulement; rien de concluant ne peut en découler. Le test du khi deux est indicatif du rapport entre le pourcentage de réussites dans la perception d'un énoncé et ses caractéristiques prosodiques. Le coefficient de Pearson mesure le degré de corrélation entre le pourcentage de réussites et les configurations tonales de chacun des domaines significatifs: le *groupe intonatif avant la transition*, la transition et le profil global de l'énoncé.

Information prosodique et/ou segmentale non manipulée

L'information prosodique et/ou segmentale non manipulée est examinée à partir de quatre expériences différentes. L'Expérience 2 consiste à associer une mélodie à une transcription; la nécessité de comprendre la notion linguistique des catégories de discours rapporté est donc éliminée. L'Expérience 3 fait appel à la compréhension de la notion linguistique, puisque la tâche consiste à distinguer si la mélodie est celle d'un /discours direct/ ou d'un discours indirect. L'Expérience 4 fait aussi appel à la compréhension de la notion linguistique, mais cette fois il s'agit d'associer un énoncé monotone à la notion de

discours direct ou de discours indirect. L'Expérience 5 présente l'énoncé dans son intégralité, à savoir information prosodique et segmentale, et utilise le type de test-notion. Ces quatre expériences de perception qui séparent, puis réunissent, information prosodique et information segmentale sont construites à partir des hypothèses suivantes.

(i) Hypothèses

1. Compte tenu que les discours rapportés comparés dans cette partie de l'étude ont été sélectionnés en fonction de leurs caractéristiques prosodiques, typiques ou atypiques,
l'hypothèse suivante est posée : (i) le pourcentage de réussites des énoncés dits prosodiquement typiques serait supérieur à celui des énoncés dits prosodiquement atypiques; (ii) le pourcentage de réussites des /discours directs/ typiques serait comparable à celui des discours indirects typiques.

2. Compte tenu que l'analyse de la production montre que c'est le domaine de la transition qui représente le plus significativement le type de discours (cf. *Figure 3.3*) et que l'analyse de la perception (corpus original) montre que la transition est le domaine qui rend compte de la plus large part des énoncés perçus (cf. section 4.3.1),
l'hypothèse suivante est posée : le domaine de la transition serait déterminant dans la perception des énoncés :
   a) une **transition montante** serait favorable à la perception du /discours direct/;
   b) une **transition descendante** serait favorable à la perception du discours indirect.
   c) Compte tenu que la transition plate est commune aux deux grandes catégories de discours rapporté (cf. section 3.2.1) et que par ailleurs la pause silencieuse est spécifique au /discours direct/ (cf. *Tableau 3.4*),
   une **transition plate** accompagnée d'une pause silencieuse serait favorable à la perception du /discours

direct/ alors qu'une transition plate non accompagnée d'une pause silencieuse sera favorable à la perception du discours indirect.

3. Compte tenu que des études sur la perception de l'accent (dont celles de LIEBERMAN:1965; ALLEN: 1972; MCDOWALL:1974; PARADIS et DESHAIES:1990; MERTENS:1991) ont montré que le pourcentage de réussites de sujets naïfs et de sujets experts était comparable,
l'hypothèse suivante est posée : dans la perception des types de discours rapporté, le pourcentage de réussites des sujets naïfs serait comparable à celui des sujets experts.

4. Compte tenu que l'absence du verbe *dire* (segmental et syntaxique) ne semble pas altérer le pourcentage de réussites alors que le mouvement de F_0 (prosodique) à la transition joue un rôle dans la perception, lorsqu'il n'y a pas de pause silencieuse (cf. *Figure 4.1*),
l'hypothèse suivante est posée : l'absence de l'information prosodique tendancielle devrait provoquer une diminution plus importante dans le pourcentage de réussites que l'absence de l'information segmentale, du moins en /discours direct/.

(ii) Expérience 2 : information prosodique test-transcription

L'objectif est de vérifier l'impact des connaissances en phonétique sur le pourcentage de réussites. La *Figure 4.4* montre les résultats.

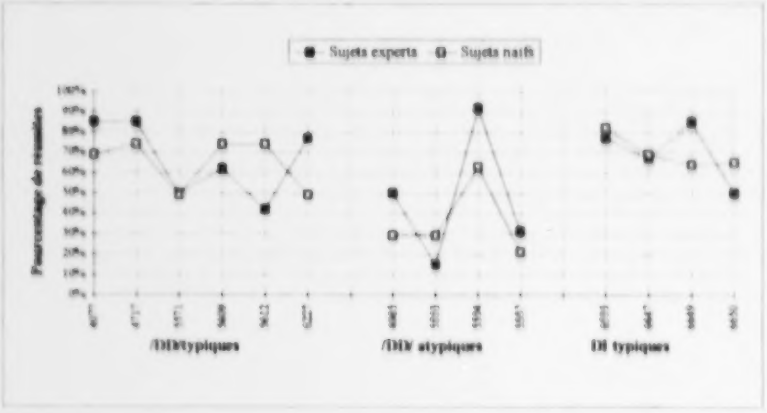


Figure 4.4 Pourcentage de réussites dans l'association des **mélodies originales** à la **transcription** d'un même énoncé en /discours direct/ (/DD/) ou en discours indirect (DI). (Expérience 2. Sujets naïfs, Groupe 3 et sujets experts, Groupe 1ex.)

1) Pour les sujets naïfs, comme pour les sujets experts, il y a relation entre le pourcentage de réussites et les caractéristiques prosodiques d'une mélodie ($p < 0,001$ pour les sujets naïfs (désormais *sn*) et $p < 0,01$ pour les sujets experts (désormais *se*).
2) Il y a relation entre le pourcentage de réussites et la seule configuration tonale de la transition; ($p < 0,01$ (*sn*) et $p < 0,05$ (*se*).
3) D'autre part, s'il existe des fluctuations particulières relativement importantes entre le pourcentage de réussites des sujets naïfs et des sujets experts, dans l'ensemble, la différence n'est pas significative.

(iii) Expérience 3 : information prosodique test-notion

L'objectif est de vérifier l'impact du type de test sur le pourcentage de réussites. L'Expérience 3 utilise à peu près les mêmes mélodies que l'Expérience 2, mais avec une tâche et un protocole différents. La *Figure 4.5* présente les résultats.

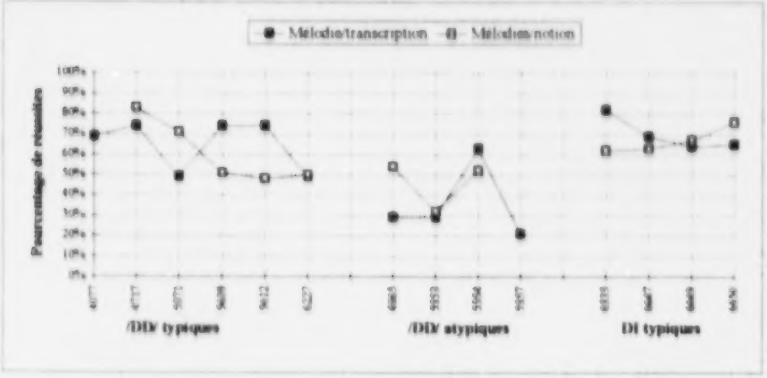


Figure 4.5 Pourcentage de réussites dans l'association des **mélodies originales** à la **transcription** en /discours direct/ (/DD/) ou en discours indirect (DI) d'un même énoncé ainsi que des **mélodies originales** à associer à la **notion** de /discours direct/ (/DD/) ou de discours indirect (DI). (Expériences 2 et 3. Sujets naïfs, Groupe 3).

1) Pour les deux types de test, il y a relation entre le pourcentage de réussites et les caractéristiques prosodiques d'une mélodie ($p < 0,001$).
2) Il y a relation entre le pourcentage de réussites et la configuration tonale de la transition ($p < 0,01$).
3) Il y a une légère différence de pourcentage moyen de réussites entre l'Expérience 2 et l'Expérience 3 : respectivement 55% et 58%. La différence est minime et se manifeste particulièrement en /discours direct/ où la variation prosodique est plus grande.

4) Dans l'Expérience 3 (à répétitions non consécutives), la typie prosodique tend à favoriser la cohérence dans le maintien des choix.

(iv) Expérience 4 : information segmentale test-notion

L'objectif est de vérifier l'impact du type d'information, prosodique (Expérience 2) par rapport à segmentale (Expérience 4). D'une part, le type de tâche est différent dans l'Expérience 2 et dans l'Expérience 4. Cependant, les résultats obtenus précédemment n'ont pas démontré que le type de tâche avait une influence sur le pourcentage de réussites. D'autre part, il est admis que les auditeurs sont très sensibles à la prosodie dès qu'il y a de l'oral (MEHTA et CUTLER:1988). Les résultats sont illustrés dans la *Figure 4.6*.

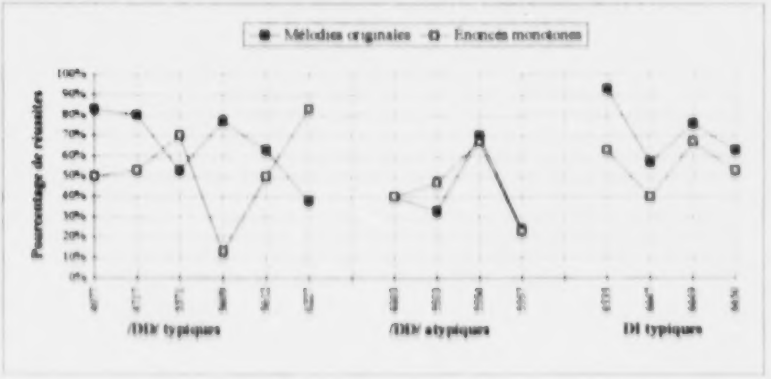


Figure 4.6 Pourcentage de réussites dans l'association des **mélodies originales** à la **transcription** d'un même énoncé en /discours direct (/DD/) ou en discours indirect (DI) ainsi que des **énoncés monotones** à associer la **notion** de /discours direct/ ou de discours indirect. (Expériences 2 et 4. Sujets naïfs, Groupe 4)

1) Tous discours confondus, la perception de la mélodie est significativement supérieure à la perception de l'énoncé monotone ($p < 0,01$). Étant donné la qualité du rendu auditif des énoncés monotones par rapport à celle des mélodies obtenues par «délexicalisation», on pourrait douter de l'entière fiabilité de ces résultats. Cependant, si l'on compare le taux de bonnes réponses des deux expériences en /discours direct/ prosodiquement typique, en /discours direct/ prosodiquement atypique et en discours indirect prosodiquement typique, on s'aperçoit qu'il n'y a qu'en /discours direct/ atypique que la variation est quasi nulle. Sans caractérisation prosodique, ce groupe est à peu près perçu de façon identique, quel que soit le type d'information supprimée; ce qui était prévisible. Ces résultats tendent à valider l'Expérience 4.

2) Cette comparaison tend à démontrer que (i) lorsqu'il y a caractérisation prosodique, la suppression de cette dernière influe significativement sur la baisse du taux de réussites; (ii) la suppression de l'information prosodique diminue davantage le pourcentage de réussites que la suppression de l'information segmentale, du moins en ce qui concerne la distinction discours direct/discours indirect.

(v) Expérience 5 : information prosodique et segmentale test-notion

L'objectif est de vérifier dans quelle mesure la remise en place de l'information intégrale, prosodique et segmentale, peut augmenter le pourcentage de réussites (par rapport à l'Expérience 3). La *Figure 4.7* montre les résultats.

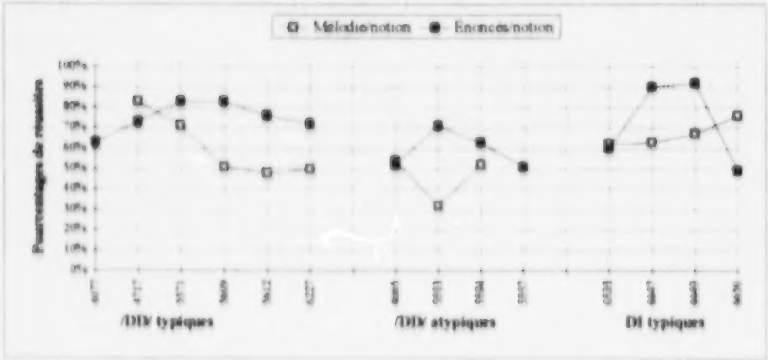


Figure 4.7 Pourcentage de réussites des **mélodies originales** ainsi que des **énoncés originaux** à associer à la **notion** de/ discours direct/ (/DD/) ou de discours indirect (DI). (Expériences 3 et 5. Sujets naïfs, Groupe 3).

1) Tel qu'attendu, il apparaît, dans l'ensemble, que l'information segmentale augmente significativement le pourcentage de réussites ($p < 0,001$).
2) Il y a cependant une relation entre le pourcentage de réussites et les caractéristiques prosodiques aussi bien en situation naturelle d'énoncé intégral ($p < 0,001$) qu'en situation artificielle de la seule mélodie ($p < 0,001$). En revanche, l'avantage de la typie du discours indirect par rapport à celle du /discours direct/, bien nette dans l'Expérience 3 ($p < 0,01$), ne l'est plus dans l'Expérience 5.
3) La transition demeure le domaine le plus corrélé au pourcentage de réussites ($p < 0,01$ dans l'Expérience 3, $p < 0,05$ dans l'Expérience 5).
4) Dans l'Expérience 5 (à répétitions non consécutives), le degré de cohérence dans le maintien des choix demeure fonction de la typie prosodique, même lorsque le contenu segmental est présent.

(vi) Comparaison entre les situations de perception

La comparaison du pourcentage moyen de perception pour les 4 expériences (*Figure 4.8*), pour chaque stimulus illustre que pour presque tous les stimuli, le taux moyen de perception des types de discours rapporté est plus élevé lorsque les caractéristiques prosodiques sont dites typiques.

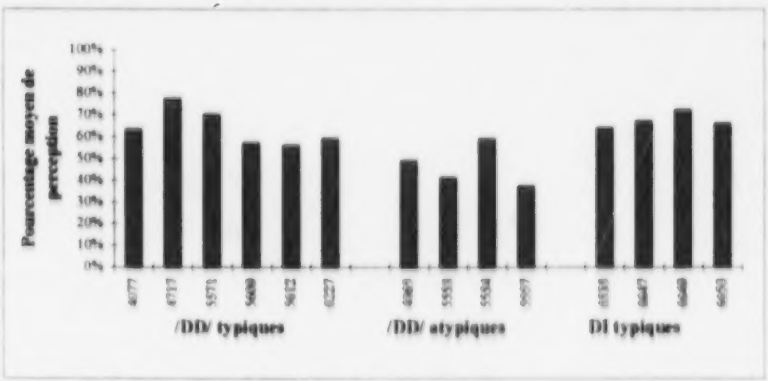


Figure 4.8 Pourcentage moyen de perception pour chaque stimulus. (Expériences 2, 3, 4 et 5. Sujets naïfs et experts).

Dans les discours rapportés porteurs de caractéristiques prosodiques dites typiques, le taux moyen de perception varie entre 56% et 77% tandis que dans les discours rapportés porteurs de caractéristiques prosodiques dites atypiques, le taux moyen oscille entre 37% et 49%, sauf pour le stimulus 5554 (59%), qui se compare au pourcentage de réussites d'un discours rapporté dit prosodiquement typique. Par ailleurs, les résultats du *Tableau 4.2* regroupent les stimuli par catégorie prosodique de discours rapporté pour les cinq situations de perception.

Tableau 4.2 Pourcentage **moyen** de perception pour chaque **situation** de perception et pour chaque **catégorie** prosodique de discours rapporté. (Expériences 2 , 3, 4 et 5. Sujets naïfs et experts).

EXPÉRIENCES	/DD/ typiques	/DD/ atypiques	DI typiques
Exp. 2 (mélodies/transcription/ naïfs)	63%	35%	68%
Exp. 2 (mélodies/transcription/experts)	65%	46%	67%
Exp. 3 (mélodies/notion/naïfs)	60%	46%	67%
Exp. 4 (énoncés monotones/notion/naïfs)	53%	44%	56%
Exp. 5 (énoncés/notion/naïfs)	75%	59%	72%
Pourcentage moyen	**63%**	**46%**	**66%**

La comparaison des pourcentages moyens de perception montre que (i) les taux moyens de perception sont comparables dans les trois situations de perception où seules les mélodies sont entendues (Expérience 2, sujets naïfs et experts et Expérience 3) : 63%, 65% et 60% en /DD/ typique, 35%, 46% et 46% en /DD/ atypique et 68%, 67% et 67% en DI typique; (ii) l'Expérience 4 ne présente pas de différence importante entre les trois sous-catégories prosodiques, puisqu'elles y sont annulées; (iii) si le pourcentage de réussites moyen augmente lorsque le prosodique s'ajoute au segmental (Expérience 5), le pourcentage de réussites demeure supérieur, lorsque les caractéristiques prosodiques sont dites typiques : 75% et 72% par rapport à 59%; (iv) le pourcentage moyen de réussites du DI typique est tout à fait comparable à celui du /DD/ typique : 66% par rapport à 63%.

Le rapport entre typie/atypie prosodique et pourcentage de réussites a préalablement été statistiquement démontré pour chacune des situations de perception; elle n'est que confirmée pour l'ensemble des cinq (5) situations de perception ($p < 0,001$).

(vii) En somme

Les expériences faites à partir de l'information prosodique et/ou segmentale non manipulée montrent que, pour les énoncés traités (dont le nombre est cependant limité), les hypothèses sont confirmées :

1. a) Le pourcentage de réussites des énoncés est toujours significativement supérieur lorsqu'un stimulus est porteur des caractéristiques prosodiques dites typiques (hypothèse 1 (i).
   b) Le pourcentage de réussites des /discours directs/ typiques n'est pas significativement différent de celui des discours indirects typiques (hypothèse 1 (ii)
2. La corrélation entre la configuration tonale de la transition et le pourcentage de réussites est toujours significative (hypothèse 2).
3. Le pourcentage de réussites des sujets naïfs n'est pas significativement différent de celui des sujets experts (hypothèse 3).
4. Le type de test-notion donne des résultats comparables à ceux du test-transcription. Il apparaît que le type de tâche n'a pas d'effet significatif sur les résultats. Aucune hypothèse n'avait été faite à ce propos.
5. L'information prosodique fournit, dans le cas présent, un pourcentage de réussites significativement supérieur à celui de l'information segmentale seule. L'information intégrale permet de montrer que, même dans ce cas, les énoncés prosodiquement atypiques sont moins bien perçus que les énoncés prosodiquement typiques (le pourcentage de réussites est de 59% (DD atypique) par rapport à 75% (DD typique) et 72% (DI typique) (hypothèse 4).

Cependant, pour réel que soit l'effet de la caractérisation prosodique du discours rapporté dans la distinction perceptuelle /discours direct/ discours indirect, la présence de ces caractéristiques prosodiques est loin d'être obligatoire. Un énoncé de discours rapporté est produit avec des caractéristiques prosodiques tendancielles environ 6 fois sur 10; c'est aussi environ 6 fois sur 10 qu'il est perçu comme

appartenant à la catégorie syntaxico-discursive appropriée lorsqu'il est présenté sous sa seule forme mélodique et 7 fois sur 10 lorsque la valeur segmentale s'ajoute à cette forme mélodique. On peut dire que, généralement, lorsqu'un discours rapporté est caractérisé prosodiquement, il est perçu dans la catégorie syntaxique appropriée. Toutefois, il arrive que des énoncés tout à fait prosodiquement caractéristiques d'un type de discours soient mal perçus et vice-versa. Il y a une zone «d'ombre» où la prosodie semble en vacances sans que les locuteurs ni les interlocuteurs aient à s'en plaindre. Mais s'agit-il d'une zone d'ombre inhérente au système linguistique ou plutôt de «finesses» qui ont échappé au filet d'analyse utilisé ? Le filet plus serré de la synthèse sera maintenant utilisé. Dans un patron intonatif donné, quel sera l'effet perceptuel (i) de la modification configurationnelle des domaines, principalement de la transition; (ii) de l'ajout ou de la suppression de la pause silencieuse; (iii) de l'inversion des caractéristiques prosodiques ?

Information prosodique manipulée

L'analyse de la perception d'un matériau original spontané a permis de confirmer que, généralement, le pourcentage de réussites des stimuli dits prosodiquement typiques est supérieur à celui des stimuli dits prosodiquement atypiques. Étant donné la corrélation significative, à répétition, entre le pourcentage de réussites et la transition, ce domaine est l'objet de la plupart des manipulations : augmentation ou diminution de F_0 à la transition, ajout ou suppression de la pause silencieuse à la transition. Les manipulations sont faites séparément :

a) augmentation ou diminution de la fréquence fondamentale à la transition;

b) ajout ou suppression de la pause silencieuse à la transition;

puis conjointement,

c) inversion des tendances fréquencielles et pausales.

Les trois expériences qui portent sur l'information prosodique manipulée sont construites à partir des mélodies à associer à la transcription d'un énoncé en discours direct ou en discours indirect (Expériences 6, 7 et 8).

(i) Hypothèses

Les hypothèses suivantes sont maintenant posées :

Première manipulation : augmentation ou diminution de F_0 à la transition

1. Compte tenu que la corrélation entre la transition et le pourcentage de réussites est forte,

   l'hypothèse suivante est posée : une augmentation de F_0 à la transition devrait (i) améliorer le pourcentage de réussites en /discours direct/ atypique et (ii) le diminuer en discours indirect dont la transition est tendanciellement descendante.

Deuxième manipulation : ajout ou suppression de la pause silencieuse à la transition

2. Compte tenu des résultats obtenus dans la perception du corpus général, à savoir que la présence d'une pause silencieuse à la transition favorise la perception du /discours direct/, indépendamment de la valeur d'écart à la transition (cf. *Figure 4.1*),
   pour le /discours direct/, l'hypothèse suivante est posée : (i) dans les /DD/ déjà typiques par leur transition montante, les modifications pausales auraient peu d'effet sur le pourcentage de réussites; (ii) dans les /DD/atypiques, les modifications pausales auraient un effet sur le pourcentage de réussites.

3. Compte tenu que la pause silencieuse à la transition est exclusive au /discours direct/, **pour le discours indirect, l'hypothèse suivante est posée : l'ajout d'une pause diminuerait le pourcentage de réussites.**

Troisième manipulation : inversion des caractéristiques prosodiques d'un type de discours

4. Compte tenu des caractéristiques prosodiques tendancielles

observées dans l'analyse de la production,
pour le /discours direct/, l'hypothèse suivante est posée : le «collage» d'une transition descendante (inférieure à -2 DT) dans les /DD/ dits typiques ainsi que la suppression de la pause silencieuse, s'il y a lieu, augmenteraient le pourcentage de réussites de ces discours dans la catégorie du DI;
pour le discours indirect, l'hypothèse suivante est posée : inversement, le «collage» d'une transition montante (supérieure à 2 DT) dans les DI dits typiques ainsi que l'ajout d'une pause silencieuse à la transition devraient augmenter le pourcentage de réussites de ces discours dans la catégorie du /DD/.

(ii) Expérience 6 : augmentation de F_0 à la transition

La manipulation consiste à vérifier l'importance d'une hausse de fréquence à la transition dans la perception des /discours directs/ atypiques et des discours indirects.

Les modifications varient entre 2,44 DT et 6,36 DT, de manière à garder le plus de naturel possible à la mélodie. Un exemple de modification est illustré à la *Figure 4.9* (le trait double signale la transition; le trait pointillé indique le début du *groupe intonatif avant la transition*). Dans la fenêtre du haut de la *Figure 4.9*, on peut observer la mélodie originale : *groupe intonatif avant la transition* plat, transition descendante et profil global descendant; la fenêtre du bas illustre les modifications effectuées : augmentation de la valeur d'écart à la transition (de -2,68 DT à 3,61 DT).

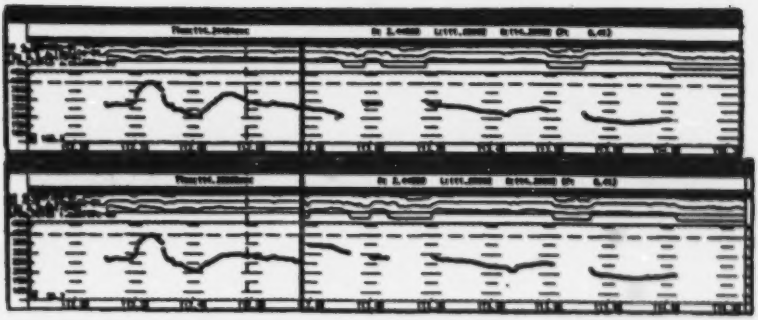


Figure 4.9 Courbe mélodique originale (fenêtre du haut) et modifiée (fenêtre du bas) de *C'est pour ça j'dis moi : «Il-y-a pas de femme heureuse comme moi.»* (5553-108). (Expérience 6).

La *Figure 4.10* permet d'illustrer l'effet de la manipulation de F_0 à la transition puisqu'elle compare le pourcentage de réussites de la mélodie originale (Expérience 2) et celui de la mélodie modifiée (Expérience 6) pour un même groupe.

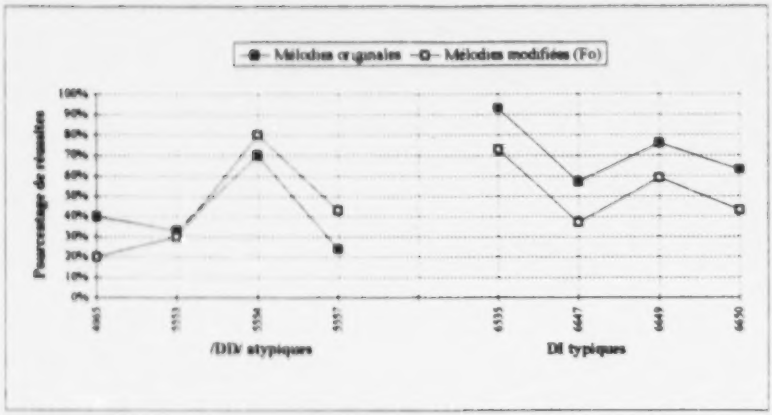


Figure 4.10 Pourcentage de réussites dans l'association des **mélodies originales** ainsi que des **mélodies modifiées (F_0 montante)** à la **transcription** en discours direct (/DD/) ou en discours indirect (DI) d'un même énoncé. (Expériences 2 et 6. Sujets naïfs, Groupe 4).

Les résultats montrent que l'augmentation de F_0 a un effet significatif sur l'ensemble des stimuli manipulés ($p < 0,05$); en fait, cet effet est particulièrement lié au discours indirect ($p < 0,01$). En /discours direct/, la seule hausse de fréquence à la transition ne permet pas d'augmenter significativement le pourcentage de réussites.

Compte tenu que (i) la reconnaissance du /discours direct/, présumée «atypique» du point de vue fréquenciel n'est pas sensiblement améliorée alors que la reconnaissance du discours indirect s'en trouve significativement diminuée, il est déduit que la seule montée de F_0 ne suffit pas à caractériser le /discours direct/.

166

(iii) Expérience 7: ajout ou suppression de la pause silencieuse

La manipulation vise à vérifier l'importance de la pause à la transition dans la distinction /discours direct/discours indirect. Les pauses silencieuses des trois (3) mélodies originales sont supprimées; des pauses silencieuses d'une durée de 250 ms sont ajoutées aux onze (11) autres mélodies. La *Figure 4.11* illustre un exemple de modification pausale. La fenêtre du haut présente la mélodie originale, avec pause, et la fenêtre du bas, la mélodie manipulée, sans pause.

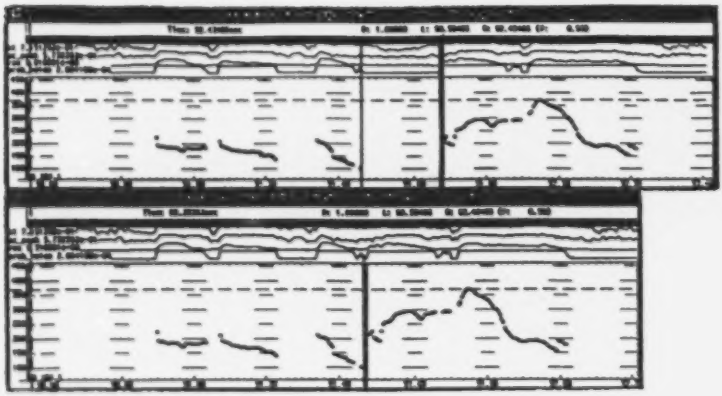


Figure 4.11 Courbe mélodique originale et modifiée de *Fait-que je lui (y) disais # : «Reste là là.»* (4717-59). (Expérience 7).

La *Figure 4.12* illustre l'effet de la modification pausale en présentant les pourcentage de réussites pour la mélodie originale (Expérience 2) et pour la mélodie modifiée (Expérience 7).

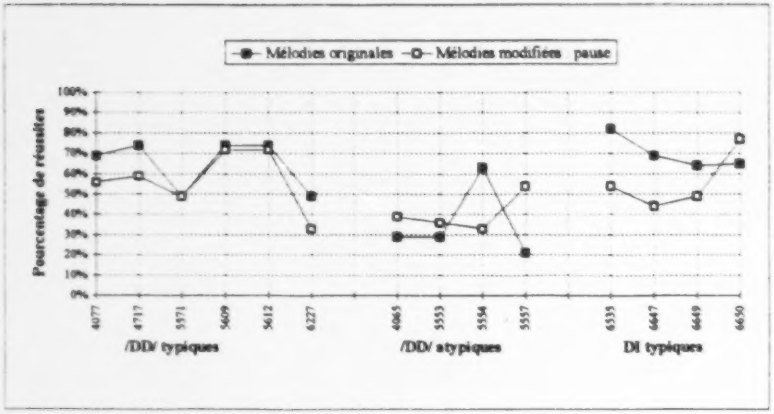


Figure 4.12 Pourcentage de réussites dans l'association des **mélodies originales** ainsi que des **mélodies modifiées (pause)** à la **transcription** en discours direct ou en discours indirect d'un même énoncé. (Expérience 2. Sujets naïfs, Groupe 3. Expérience 7. Sujets naïfs, Groupes 5 et 6).

Il apparaît que pour l'ensemble des discours, l'ajout ou la suppression de la pause silencieuse n'a pas vraiment d'effet significatif sur le pourcentage de réussites ($p = 0,064$). En /discours direct/ typique, dans l'ensemble, les modifications pausales n'ont pas d'effet significatif sur le pourcentage de réussites. En /discours direct/ atypique, là où une amélioration du pourcentage de réussites était attendu, l'effet n'est pas significatif dans l'ensemble.

On se souvient que le modèle de régressions multiples construit à partir des résultats obtenus pour la perception du corpus général (à partir de l'information prosodique et segmentale) prédit que la présence d'une pause silencieuse favorisera la perception en /discours direct/ dans 65% des cas, indépendamment de la valeur d'écart à la transition. Cette prédiction est loin de se vérifier ici. Il est vrai que l'information segmentale est absente et que le nombre d'énoncés en /discours direct/

est fort limité (10). L'Expérience 7, comme la précédente, semble montrer qu'un élément prosodique donné ne peut jouer qu'un rôle limité dans la perception et qu'en /discours direct/ spontané, les combinaisons prosodiques sont assez fluctuantes.

En revanche, en ce qui concerne les DI typiques, dans l'ensemble, l'ajout pausal diminue significativement le pourcentage de réussites dans la perception du discours indirect, et de façon plus marquée chez les sujets experts; ($p < 0,05$ pour les *sn*, $p < 0,01$ pour les *se*).

(iv) Expérience 8 : inversion des tendances prosodiques (F_0 et pause)

La manipulation consiste à vérifier dans quelle mesure l'inversion des caractéristiques prosodiques peut engendrer une augmentation du pourcentage de réussites dans la perception du discours de la catégorie syntaxique opposée.

Les modifications fréquencielles varient entre -2,61 DT et -5,51 DT pour les /DD/ typiques et entre 4,10 DT et 6,36 DT pour les DI typiques. Comme pour l'Expérience 6, l'ampleur de l'écart a été déterminée de manière à conserver le plus possible à la mélodie un «son» naturel; tous les verbes introducteurs ont été abaissés, ou rehaussés. Dans certaines mélodies, le *groupe intonatif avant la transition* a aussi été modifié dans le sens opposé à la modification sur le verbe introducteur et dans un cas de /DD/ typique, le profil global de l'énoncé a été abaissé.

Les modifications pausales consistent à supprimer la pause dans les /DD/ typiques qui en contiennent et à ajouter une pause dans les DI typiques.

La *Figure 4.13* présente un exemple des modifications fréquencielles effectuées. La fenêtre du haut présente la mélodie originale avec *groupe intonatif avant la transition* plat, transition

montante et profil global montant. La fenêtre du bas présente la mélodie modifiée : *groupe intonatif avant la transition* montant, transition descendante et profil global descendant.

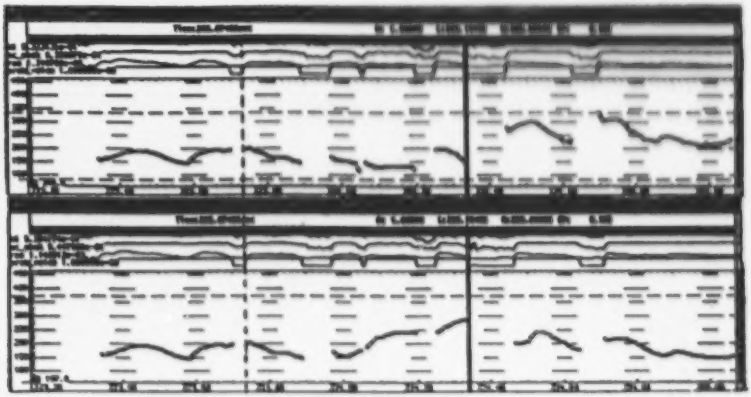


Figure 4.13 Courbe mélodique originale et modifiée de *Les autres m'encourageaient : «Cours cours Bina.»* (5609-108) (Expérience 8).

La *Figure 4.14* montre l'effet de l'inversion des caractéristiques fréquencielles et pausales sur le pourcentage de réussites.

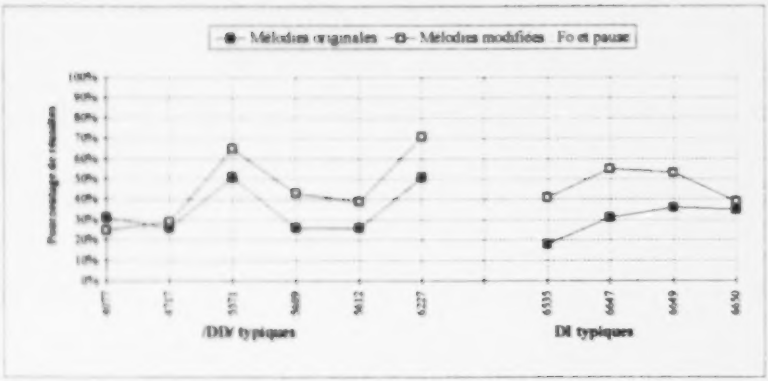


Figure 4.14 Pourcentage de réussites dans l'association des **mélodies originales** ainsi que des **mélodies modifiées (F_0 et pause)** à la **transcription** en /discours direct/ (/DD/) ou en discours indirect (DI) d'un même énoncé dans la catégorie opposée. (Expériences 2. Sujets naïfs, Groupe 3. Expérience 8. Sujets naïfs, Groupes 6 et 7).

On constate que les modifications ont généralement atteint leur but : augmenter le pourcentage de réussites dans la perception d'un type discours de la catégorie opposée. Qu'on considère ensemble les deux catégories de discours rapporté ($p < 0,001$) ou qu'on les considère séparément (/DD/, $p < 0,05$; DI, $p < 0,01$), l'effet des modifications est significatif. La corrélation entre le pourcentage de réussites et la configuration tonale des domaines n'est cependant pas significative alors qu'elle l'était à la transition, dans l'Expérience 2 ($p < 0,01$). Les résultats sont un peu moins marqués chez les sujets experts. Il reste que les pourcentages de réussites ont généralement été déplacés dans la direction voulue.

(v) En somme

Ces expériences avec de l'information prosodique manipulée montrent que, pour les données traitées :

1) La seule augmentation de F_0 à la transition, contrairement à

l'hypothèse 1 (i), ne suffit pas à augmenter le pourcentage de réussites en /discours direct/;

2) la perception du discours indirect se trouve en revanche passablement diminuée lorsqu'il y augmentation de F_0 à la transition (hypothèse 1 (ii).

3) Tel qu'attendu en /discours direct/ typique, les modifications pausales n'ont pas d'effet significatif sur le pourcentage de réussites (hypothèse 2 (i). En revanche, contrairement à ce qui était attendu en /discours direct/ atypique, l'ajout de la pause silencieuse n'a pas d'effet significatif sur le pourcentage de réussites (hypothèse 2 (ii).

4) Conformément à l'hypothèse, la présence de la pause silencieuse à la transition en discours indirect contribue par ailleurs, comme l'augmentation de F_0, à diminuer le pourcentage de réussites, particulièrement chez les sujets experts (hypothèse 3).

5) L'inversion des tendances prosodiques provoque une augmentation significative du pourcentage de réussites dans la catégorie opposée, surtout en discours indirect. En /discours direct/, l'effet est sensible chez les sujets naïfs, mais à peine chez les sujets experts. L'augmentation n'entraîne pas pour autant une «transperception» (≥ 60%) (hypothèse 4).

Conclusion

D'une part, l'analyse de la perception d'une information prosodique **originale** a montré que, règle générale, lorsqu'une information prosodique est caractéristique d'un type de discours, elle favorise la perception du type de discours approprié. L'analyse de la perception d'une information prosodique **manipulée** démontre d'autre part que (i) l'ajout des principales caractéristiques considérées typiques du /discours direct/, transition montante ou pause silencieuse, à des mélodies originales du /discours direct/, dites prosodiquement atypiques, ne provoque pas des changements considérables dans le pourcentage de réussites, alors que (ii) l'insertion de ces mêmes caractéristiques dites typiques du /discours direct/ dans une mélodie de

discours indirect a, en revanche, pour effet de diminuer de façon importante la perception du discours indirect. Il en ressort que la stabilité prosodique du discours indirect découle de la stabilité syntaxique de sa structure, structure par ailleurs tout à fait comparable à celle d'un énoncé déclaratif quelconque. Et bien sûr, cette stabilité prosodique favorise la perception du discours indirect.

En revanche, lorsque la stabilité syntaxique est plus relative, comme c'est le cas en /discours direct/, la stabilité prosodique semble davantage reliée à la valeur discursive. Dans le sous-corpus retenu, les 6 /discours directs/ prosodiquement typiques qui ont été les mieux perçus correspondaient au type d'emploi de la *reproduction* (valeur narrative la plus forte) alors que les 4 /discours directs/ prosodiquement atypiques, moins bien perçus, correspondaient à divers types d'emploi discursif. Il a été montré précédemment (cf. section 3.3.1) que la transition montante était fonction d'une valeur narrative forte. Mais alors comment expliquer que l'augmentation de F_0 à la transition[15] n'ait pas donné les effets escomptés sur le pourcentage de réussites dans la perception du /discours direct/ ?

L'introduction de la tendance prosodique la plus corrélée à la valeur discursive du /discours direct/ (augmentation de F_0 à la transition) n'a pas désambiguïsé complètement la perception d'un /discours direct/ dit prosodiquement atypique. La valeur narrative n'englobe probablement pas toutes les valeurs discursives liées à l'utilisation du /discours direct/. Par ricochet, les marques prosodiques sont sans doute plus diversifiées que celles ici étudiées. Pour bien réel que soit le lien entre la valeur narrative, un élément discursif, et la caractérisation fréquencielle, un élément prosodique, il faut pourtant admettre qu'il s'agit là d'un élément discursif et d'un élément

[15] Rappelons que du point de vue discursif, aucune corrélation n'a été observée entre la valeur discursive du /discours direct/ et la pause silencieuse (cf. section 3.3.2).

prosodique parmi d'autres. Il faut ici élargir la notion de valeur discursive de même que celle des moyens utilisés dans l'identification du /discours direct/ :

(i) La valeur narrative ne paraît pas recouvrir toutes les valeurs discursives rattachées au /discours direct/. Pour une même valeur narrative, le sentiment, l'émotion du rapporteur en regard des paroles citées peut varier (joie, mépris, timidité, etc); il en est de même pour l'attitude (*accord* ou *désaccord* par exemple). Il est alors possible que des moyens linguistiques, tels le lexique ou la syntaxe, ainsi que des moyens extra-linguistiques, tels la situation ou les gestes, puissent s'ajouter, ou suppléer, à la prosodie.

(ii) Cette diversité de valeurs discursives pourrait aussi être représentée par des marques prosodiques supplémentaires, comme l'intensité, et/ou des marques paralinguistiques, comme la qualité de la voix par exemple.

L'analyse de la production avait montré que la prosodie du /discours direct/ était corrélée à la valeur discursive, notamment à celle du type d'emploi de la *reproduction*. L'analyse de la perception montre que la valeur narrative ne peut supporter à elle seule toute la variabilité discursive du /discours direct/, et par là justifier l'ensemble des variations prosodiques.

La prochaine section s'intéresse justement à la fonction «attitudinale» de l'intonation dans le discours rapporté. L'objectif n'est pas de mettre en évidence les moyens linguistiques en général, ni même de dégager les paramètres prosodiques représentatifs d'une attitude, mais simplement de s'assurer que la mélodie du discours rapporté peut effectivement entraîner la perception d'une attitude spécifique et si tel est le cas, d'examiner dans quelle mesure la fréquence et le débit peuvent en être des marqueurs.

214

4.7.2 PERCEPTION DE LA DIFFÉRENCE D'ATTITUDE

La grande majorité des recherches en prosodie s'intéresse à la fonction syntaxique de celle-ci, bien qu'un certain nombre d'études se soit également intéressé à une fonction informative (entre autres HALLIDAY:1967, 1970, 1976; CRYSTAL:1969, 1975; LADD:1980; CRUTTENDEN:1986), textuelle (tels FOX:1973; BRAZIL *et al.*:1980; YULE:1980; COUPER-KUHLEN:1986; TENCH:1990) de même qu'à celle qui nous occupe ici, une différence d'attitude (dont LIEBERMAN et MICHAELS:1962; ULDALL:1960, 1964; CRYSTAL:1969; O'CONNOR et ARNOLD:1973; LADD *et al.*:1984; GUAÏTELLA:1991; VROOMEN *et al.*:1993; ABADJIEVA *et al.*:1993; MOZZICONACCI:1995)[16]. Il est raisonnable de croire que dans tout énoncé, la prosodie joue simultanément, ou alternativement, à divers degrés, ces différents rôles. La question ici n'est pas de savoir si la prosodie joue un rôle «attitudinal» —c'est probable— mais plutôt de voir dans quelle mesure elle joue ce rôle et dans quelle mesure les paramètres prosodiques étudiés, fréquence et débit, sont des éléments actifs sur le plan «attitudinal». Dans la présente étude, l'attitude assez générale d'*accord* ou de *désaccord* du locuteur par rapport aux paroles citées est retenue. Cette attitude d'*accord* ou de *désaccord* peut très bien être empreinte d'émotion. On peut ici tenter de distinguer *émotion* et *attitude* en partant du point de vue que l'*émotion* relève d'une activité physiologique (non linguistique) alors que l'*attitude* relève d'une activité cognitive (linguistique)[17] (cf. COUPER-KUHLEN 1986:173-174).

L'émotion désigne plutôt ce qui est involontaire (la joie, l'ennui, la colère, etc.); l'attitude désigne davantage ce qui est volontaire (la compréhension, la politesse, l'approbation, etc.). En principe donc, la

---

16 Pour une revue exhaustive des travaux sur les diverses fonctions de l'intonation, on consultera avec profit COUPER-KUHLEN (1986) et TENCH (1990).

17 "An angry person does not raise his voice in English or in German but simply in anger" (LIEB 1980:34 cité dans COUPER-KUHLEN 1986:173).

manifestation de l'émotion, involontaire, serait davantage tributaire de marques paralinguistiques (hésitation, rire, qualité de la voix, etc.) et extralinguistiques (hochement de la tête, élévation des sourcils, mouvement des mains, etc.). La manifestation d'une attitude, volontaire, serait pour sa part plus liée à des moyens linguistiques : le lexique, la construction syntaxique et bien sûr la prosodie, qui nous intéresse plus spécifiquement ici. C'est, globalement, la position de SILVERMAN et *al.* (1983). Ce qui ne veut pas dire que les moyens linguistiques, paralinguistiques et extralinguistiques soient associés de façon univoque à un seul type d'expression.

Ainsi, une étude de VROOMEN *et al.*(1993) met en relation des émotions (la neutralité, la joie, l'ennui, la colère, la tristesse, la peur et l'indignation) avec les paramètres linguistiques de débit, d'intonation et de débit-intonation. Il en ressort que les manipulations prosodiques effectuées permettent de confirmer des rapports privilégiés entre des moyens linguistiques, en l'occurrence prosodiques, et une émotion. Par ailleurs, les fondements éthologiques, physiologiques et psychologiques (non linguistiques) de l'émotion ont récemment été illustrés dans une étude comparative entre le coréen et le français. En effet, CHUNG (1995) montre que les deux langues présentent des caractéristiques similaires. La qualité de la voix, caractéristique paralinguistique, permettrait de distinguer les émotions positives (telles la joie et la tendresse) des émotions négatives (telles la colère et la tristesse) alors que le registre, le mouvement tonal, le débit et l'intensité, caractéristiques linguistiques, permettraient de distinguer les émotions actives (la joie et la colère) des émotions passives (la tendresse et la tristesse).

Bien entendu, l'attitude *accord* ou *désaccord* avec les paroles citées peut être porteuse d'émotions différentes. Ainsi le locuteur peut-il être en *accord* avec les paroles citées et heureux, triste ou résigné; ou en *désaccord* et ennuyé, effrayé ou fâché, etc. À l'évidence, l'attitude d'*accord* entraîne davantage des émotions positives tandis que l'attitude de *désaccord* entraîne plutôt des émotions négatives, types d'émotion

qui seraient surtout marqués par une caractéristique paralinguistique, la qualité de la voix (CHUNG 1995:268). Il apparaît alors illusoire de vouloir détecter toute la gamme d'émotions[18] de l'attitude *accord/désaccord* à partir des paramètres prosodiques seulement, ici la fréquence et le débit.

Il faut donc oublier les nuances et s'en tenir à vérifier l'importance de la prosodie par rapport à l'importance du segmental dans la perception de l'attitude *accord/désaccord*. Enfin, rappelons que l'étude est basée sur des extraits de corpus spontané, ce qui est bien différent d'énoncés lus, destinés spécifiquement à transmettre une valeur expressive (comme c'est le cas dans l'étude de MOZZICONACCI:1995 par exemple). Partir d'un corpus spontané rend la systématisation moins probable, mais la représentation, aussi éclatée soit-elle, devrait être un reflet plus fidèle de ce qui se passe dans l'environnement conversationnel naturel.

Information segmentale et/ou prosodique

Le discours indirect n'est pas exempt de l'attitude du rapporteur en regard des paroles citées, mais le marquage de celle-ci serait davantage associable au lexique (AUTHIER:1978; MAINGUENEAU:1986) alors qu'en /discours direct/, le marquage de l'attitude serait davantage associable à l'intonation. Pour cette raison, les expériences de perception de l'attitude ne portent que sur du /discours direct/. L'analyse de la perception «attitudinale» vise à vérifier dans quelle mesure l'attitude *accord/désaccord* du locuteur par rapport aux paroles citées est perceptible, d'abord à travers l'énoncé original, porteur d'une construction grammaticale, d'un lexique et d'une prosodie qui lui est propre (Expérience 9; cf. *Annexe B*), puis à travers la seule prosodie de ces mêmes stimuli (Expérience 10). Les Expériences 9 et 10 sont

[18] On se rappelle les propos d'AUTHIER (1978) concernant «toutes les nuances possibles [...] de l'agressivité polémique à la complicité tendre» pouvant être traduites par l'intonation (cf. section 1.2.2).

construites à partir des hypothèses suivantes.

(i) Hypothèses

1. Compte tenu que des analystes du discours (dont AUTHIER:1978) évoquent la manifestation d'une attitude du locuteur par rapport aux paroles citées,
l'hypothèse suivante est posée : le stimulus intégral permettrait de reconnaître de façon relativement nette une attitude d'*accord* ou de *désaccord*.

2. Compte tenu que la distinction *accord/désaccord* se rapproche de la distinction émotion positive/émotion négative et que selon l'étude de CHUNG (1995), cette dernière distinction repose sur la qualité de la voix,
l'hypothèse suivante est posée : la reconnaissance de l'attitude ne serait pas fonction des caractéristiques prosodiques ici retenues, la fréquence fondamentale et le débit.

3. Compte tenu qu'en /discours direct/, l'intonation est reconnue comme marqueur d'attitude (AUTHIER:1978; MAINGUENEAU:1981)
l'hypothèse suivante est posée : l'élimination du segmental diminuerait le taux de perception, mais la classification dans une attitude serait maintenue.

(ii) Expérience 9 : information segmentale et prosodique/test-transcription

L'objectif consiste à vérifier la polarisation du pourcentage de concordance autour d'une attitude d'*accord* ou de *désaccord*. La sélection des énoncés de l'Expérience 9 repose sur une valeur d'attitude pressentie par l'analyste. Cependant, contrairement aux expériences précédentes, il n'y avait pas de bonnes ou de mauvaises réponses. Il s'agissait de vérifier dans quelle mesure chacun des énoncés retenus pouvait rallier une majorité de sujets (≥ 60%) dans une attitude d'*accord* ou de *désaccord*.

L' Expérience 9 porte sur six énoncés. Il s'agit d'associer un énoncé à une transcription accompagnée de la notion explicitée d'*accord* ou de *désaccord* du locuteur par rapport aux paroles rapportées (cf. *Annexe B*). Par exemple :

(2) Ils l'mettent sur le tableau : «Pi arrangez-vous avec le tableau.» (5591-108)

- □ **(ACCORD : C'est la bonne façon de faire.)**
- □ **(DÉSACCORD : Ce n'est pas la bonne façon de faire.)**

Les résultats montrent, dans un premier temps, que l'audition de l'information intégrale permet de classer de façon assez nette chacun des énoncés dans une attitude d'*accord* ou de *désaccord* (hypothèse 1). Le pourcentage moyen de concordance pour les 6 énoncés est de 72%. La différence du pourcentage de concordance dans la perception de la notion d'*accord* (taux moyen de 70%) et de *désaccord* (taux moyen de 74%) n'est pas significative. Dans un deuxième temps, ces résultats montrent qu'aucun des paramètres prosodiques étudiés n'intervient de façon significative dans la perception de l'attitude *accord/désaccord* (hypothèse 2). Les facteurs prosodiques retenus ne pouvant être mis en rapport avec le pourcentage de concordance, il peut être déduit que des éléments linguistiques tels le lexique, la construction syntaxique ou des éléments paralinguistiques telles des caractéristiques liées à la qualité de la voix (CRYSTAL:1969; CHUNG:1995) sont des facteurs de reconnaissance de l'attitude. L'Expérience 10 élimine l'information segmentale afin d'observer la marge d'action de cette dernière.

(iii) Expérience 10 : information prosodique/test-notion

L'objectif de cette expérience est d'observer dans quelle mesure le retrait du segmental peut diminuer le pourcentage de concordance. La tâche consiste cette fois à associer la mélodie à la notion d'*accord* ou de *désaccord*.

La *Figure 4.15* permet de comparer les résultats de l'Expérience 10 à ceux de l'Expérience 9.

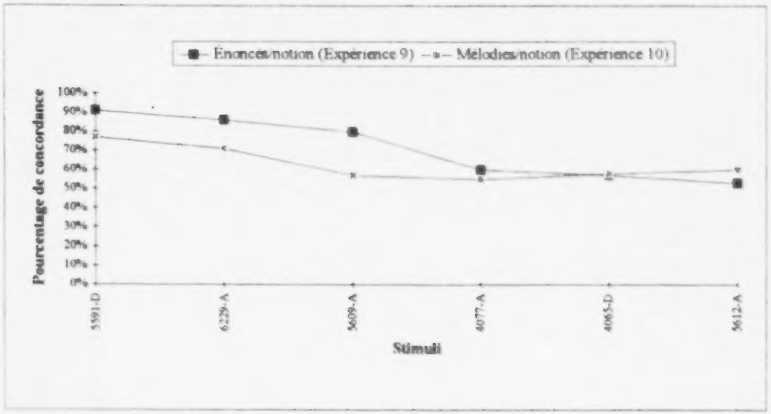


Figure 4.15 Pourcentage de concordance dans l'association des **énoncés** et des **mélodies** à l'attitude du locuteur —*accord* (A) ou *désaccord* (D)— par rapport au discours cité. (Expériences 9 et 10. Sujets naïfs, Groupes 7 et 8).

Tel qu'attendu, l'élimination du segmental a un effet diminutif sur le pourcentage de concordance dans la perception de l'attitude (le pourcentage moyen de concordance passe de 72% à 63%; la différence est significative : $p < 0,01$) (hypothèse 3). Chaque stimulus est toutefois classé selon la même attitude que lorsque l'information segmentale est présente. C'est donc dire que la prosodie joue un rôle important dans la distinction des attitudes. Par ailleurs, il semble que cette importance ne soit pas la même pour les deux attitudes ($p < 0,05$). En effet, la différence de perception entre la notion d'*accord* et de *désaccord* est ici significative. Les deux mélodies reconnues comme porteuses d'une attitude de *désaccord* sont significativement mieux perçues que celles porteuses d'une attitude d'*accord*, ce qui n'est pas le cas lorsque toute l'information linguistique est présente. À première vue, cela porte à croire que le prosodique a plus d'importance dans la reconnaissance du *désaccord* (émotion négative) que dans la reconnaissance de l'*accord* (émotion positive). Cette observation

demanderait évidemment à être confortée par un corpus beaucoup plus étendu.

Dans cette expérience, le degré de cohérence n'apparaît pas être fonction du type d'attitude, pas plus d'ailleurs que de la typie ou de l'atypie prosodique.

Les paramètres prosodiques retenus n'entretiennent pas de relation significative avec le pourcentage de concordance. Dans le même sens, on remarque que les mélodies pour lesquelles le pourcentage de concordance est le plus élevé sont des mélodies dites prosodiquement atypiques dans le présent corpus. Il semble bien que les paramètres prosodiques retenus sont davantage liés à la valeur narrative (cf. section 3.3.4) qu'à la valeur d'attitude, valeur qui serait peut-être mieux représentée par la qualité de la voix, tel que posé dans l'hypothèse 3. Certes, il faudrait des vérifications plus précises.

Dans une étude récente, MOREL (1996) pose pour sa part qu'il y a corrélation entre l'intonation et les valeurs discursives «d'adhésion» et de «désengagement» en discours direct oral spontané. Cependant, les correspondances terminologiques et méthodologiques entre l'étude de MOREL et la présente étude sont trop hasardeuses pour faire un lien entre les deux études.

(iv) En somme

Cette brève incursion du côté de la fonction «attitudinale» de la prosodie dans la reconnaissance de l'attitude d'*accord* ou de *désaccord* du locuteur par rapport aux paroles citées montre que :

1) Il y a la reconnaissance d'une attitude de façon assez nette (un taux moyen de 72%) lorsque l'information intégrale est donnée (segmentale et prosodique). Le taux de reconnaissance diminue (63%) lorsque seulement l'information prosodique est donnée, mais la classification de chaque énoncé dans une attitude est maintenue. La prosodie joue donc un rôle important dans la reconnaissance de

l'attitude, du moins dans celle d'*accord* ou de *désaccord*.

2) Cependant, les facteurs prosodiques ici étudiés, la fréquence et le débit, ne jouent pas de rôle significatif dans la reconnaissance de cette attitude.

3) D'autres facteurs prosodiques, telle l'intensité, ou paralinguistiques, comme la qualité de la voix par exemple, sont probablement mis à contribution; l'activité mimo-gestuelle également.

4.8 CONCLUSION

Au terme de cette analyse de la perception de la prosodie dans le discours rapporté, il apparaît que, généralement, en ce qui concerne la perception de la **différence syntaxique** :

1) lorsque des caractéristiques prosodiques tendancielles sont présentes, elles favorisent la reconnaissance du type de discours rapporté;

2) les marques prosodiques du discours indirect sont relativement stables : lorsque l'information prosodique originale est donnée, le pourcentage de réussites est relativement stable alors que lorsque l'information prosodique est manipulée en introduisant des caractéristiques du /discours direct/, le pourcentage de réussites diminue sensiblement;

3) les marques prosodiques du /discours direct/ sont plus fluctuantes : lorsque l'information prosodique originale est donnée, le pourcentage de réussites est variable et lorsque l'information prosodique est manipulée en renforçant les caractéristiques prosodiques, le pourcentage de réussites ne s'en trouve généralement pas sensiblement affecté, même dans l'Expérience 8 où l'effet conjoint F_0 -pause a été examiné.

182

4) D'une part, il reste que l'effet conjoint F_0 -débit n'a pas été véritablement analysé. D'autre part, la variété des motifs discursifs du /discours direct/ pourrait demander une variété de moyens plus étendue que les seuls facteurs prosodiques de F_0, de la pause silencieuse et du débit; l'intensité joue probablement un rôle non négligeable. De plus, des moyens paralinguistiques comme le rire, l'hésitation et la qualité de la voix sont certainement mis à contribution, de même que le système mimo-gestuel.

En ce qui concerne la perception de la **différence d'attitude**:

5) En /discours direct/, le rôle «attitudinal» de la prosodie apparaît aussi perceptible que le rôle grammatical. L'attitude (*accord* ou *désaccord*) rattachée à chacun des énoncés porteur de l'information intégrale est maintenue lorsque seule l'information prosodique est présentée. La perception d'une attitude apparaît assez liée à la prosodie.

6) Les facteurs prosodiques ici analysés (F_0 et débit) ne permettent cependant pas de caractériser prosodiquement l'attitude d'*accord* et de *désaccord*, possiblement liée à l'intensité ou même à la qualité de la voix. La relation entre la prosodie et les valeurs discursives d'attitude mériterait d'être approfondie.

La prosodie fait partie des paramètres linguistiques à la base de la communication. Cependant, la conversation naturelle est faite de plusieurs niveaux d'information qui paraissent s'opposer à une spécialisation étroite de chaque niveau : la constante se trouverait plutôt dans le va-et-vient continuel d'un niveau à un autre.

CONCLUSION

Des spécialistes de l'intonation tels BOLINGER, CRUTTENDEN, COUPER-KUHLEN affirment que la prosodie et la syntaxe ne s'associent que de façon bien relative; des analystes du discours, comme AUTHIER et MAINGUENEAU, spécialistes du discours rapporté, accordent une grande importance à l'intonation dans la distinction discours direct/discours indirect. Qu'en est-il dans les faits, en français québécois oral spontané ? Tel a été le point de départ de cette thèse.

D'une part, l'étude a été orientée de façon à montrer les caractéristiques prosodiques du discours rapporté en français québécois plutôt qu'à montrer l'importance de la variété dialectale dans la prosodie du discours rapporté. D'autre part, le fait de traiter des énoncés tirés de la conversation courante a été considéré comme une variable susceptible de présenter des caractéristiques prosodiques plus fluctuantes que celles des énoncés lus : les premiers sont vraisemblablement soumis à la situation du moment, éminemment variable, alors que les seconds sont soumis aux règles de la «bonne» lecture à haute voix, plutôt fixes. Enfin, la prise en compte des facteurs sociaux est apparue comme allant de soi.

Lorsque le corpus a été mis en place, l'analyse quantitative envisagée a tout de suite soulevé le problème général du quoi mesurer et du comment le mesurer ? En termes plus spécifiques, quel modèle de représentation de l'intonation était le plus adéquat pour rendre compte de l'objectif fixé : décrire l'organisation prosodique du discours

rapporté à partir d'un corpus d'oral spontané ? C'était la première d'une série de questions générales posées au début de cette recherche.

Au moment de conclure, je rappelle simplement ces questions ainsi que les réponses données et discutées en cours d'analyse afin de mettre en évidence les aspects importants de l'étude en même temps que d'évaluer dans quelle mesure les objectifs ont été atteints.

QUESTIONS POSÉES

(1) Quel est le modèle de représentation de l'intonation le plus approprié au traitement du présent corpus de français québécois oral spontané ?

En parole spontanée :

(2) Est-ce qu'il y a une distinction entre la structure syntaxique du discours direct et la structure syntaxique du discours indirect du point de vue de la fréquence fondamentale, de la pause et du débit ?

(3) En discours direct, la structure prosodique est-elle sensible à la variation syntaxique propre à l'énoncé rapporté en parole spontanée : présence ou absence du verbe introducteur (*dire*) ainsi que présence ou absence d'un «marqueur transitionnel» (pauses ou marqueurs d'attaque comme *ah*, *bon*, *ben*, etc.) ?

(4) Les caractéristiques prosodiques sont-elles fonction de la valeur discursive de l'énoncé rapporté ? Si oui, jusqu'où la finesse prosodique peut-elle aller dans le marquage des valeurs discursives : est-ce qu'il y a un marquage binaire, expressif ou neutre, ou un marquage n+aire, rattaché à la valeur narrative, plus ou moins narrative, ou encore rattaché à des valeurs «attitudinales» multiples ?

(5) Dans quelle mesure les traits prosodiques produits sont-ils fonction de facteurs sociaux tels le sexe, l'âge et le statut socioprofessionnel ?

(6) Dans quelle mesure les caractéristiques prosodiques produites sont-elles perçues ?

RÉPONSES DONNÉES

(1) *LA REPRÉSENTATION DE L'INTONATION*

Les différents modèles disponibles ne font pas consensus sur la façon de représenter l'intonation. Il importait donc d'en faire une revue critique de façon à choisir un mode de représentation approprié au type de corpus ici traité : du français, oral spontané.

Parmi les approches descriptives, plus adaptées au type de corpus traité, les modèles de GARDING, de HIRST et DI CRISTO et de MERTENS, ont déjà été appliqués au français. L'approche ici retenue est qualifiée de mixte en ce qu'elle combine une étude instrumentale au modèle de MERTENS, élaboré à partir d'une étude impressionniste. Le choix de ce modèle se justifie principalement par le fait qu'il est issu d'un corpus spontané, mais aussi par le fait que la définition précise de chacune des composantes en facilite l'application. De plus, des éléments propres à l'oral spontané, comme les troncations, les répétitions, les hésitations y trouvent leur place à l'intérieur d'un groupe intonatif, ou constituent en eux-mêmes un groupe intonatif, selon leur valeur tonale. La règle de hiérarchie des tons du système de MERTENS permet ainsi de délimiter les groupes intonatifs et de les regrouper sur une base uniquement prosodique, objectivité appréciable pour une structure aussi marquée syntaxiquement que celle du discours rapporté.

Évidemment, comme la méthode impressionniste de MERTENS a ici été remplacée par une analyse instrumentale, la façon de

186

neutraliser les variations microprosodiques contextuelles (mesure aux deux tiers de la voyelle), le point de prise de mesure, les mesures intersyllabiques et l'établissement de seuils de perception pour les intervalles mineur et majeur ainsi que pour les mouvements de F_0 demeurent discutables. Je crois cependant que, dans l'état actuel des connaissances, la représentation de l'intonation des énoncés du présent corpus sert bien la tâche fixée : l'analyse de l'organisation prosodique du discours rapporté.

(2) *LES DIFFÉRENCES PROSODIQUES ENTRE LE /DISCOURS DIRECT/ ET LE DISCOURS INDIRECT*

Globalement, le discours indirect est prosodiquement différent du discours direct. Le discours indirect est le type de discours rapporté qui présente les caractéristiques prosodiques les plus stables (environ 6 fois sur 10 dans le meilleur des cas) : montée dans la partie citante, chute dans la partie citée, jusqu'à la fin de l'énoncé, comme un énoncé neutre; pas de «marqueurs transitionnels» (pause et/ou marqueur d'attaque comme *ah, ben, coudon,* etc.) entre la partie citante et la partie citée. Il est à remarquer qu'il en est ainsi du point de vue syntaxique. Le discours indirect contient toujours un verbe de parole et il est très rare que la conjonction de subordination *que* soit absente.

D'autre part, les /discours directs/ ont des caractéristiques prosodiques plus fluctuantes (la récurrence d'une tendance est d'environ 4 fois sur 10), bien qu'elles soient généralement contrastives par rapport à celles du discours indirect : en effet, la partie citante est plus souvent descendante ou plate que montante, la transition est plus souvent montante ou plate que descendante, le profil global de l'énoncé est plus souvent montant ou plat que descendant; les «marqueurs transitionnels» sont courants entre la partie citante et la partie citée. Cette variabilité prosodique est reflétée dans la structure syntaxique : tantôt est inclus un verbe de parole, tantôt il est omis; tantôt un «marqueur transitionnel» sépare la partie citante de la partie cité, tantôt pas.

Il découle déjà de ces premières observations une remise en question quant à la force de la relation entre prosodie et syntaxe, particulièrement en /discours direct/, du moins en spontané. En effet, dans l'étude de LÉANDRI (1993), effectuée à partir d'un corpus lu, le /discours direct/ n'oscillait pas entre deux tendances fréquencielles. On est à même de constater que la parole spontanée a une mélodie plus variée, plus riche, que la parole lue. Toutefois, les éléments constitutifs de cette «richesse» sont encore flous en ce qui concerne précisément le discours rapporté.

Du point de vue du débit, le /discours direct/ ne s'oppose pas au discours indirect. C'est plutôt le discours direct sans verbe *dire* qui s'oppose au discours direct avec verbe *dire*, lui-même tout à fait semblable au discours indirect. Le débit apparaît ici plus sensible au «trou syntaxique» qu'au type de construction syntaxique (indépendante ou complétive). De toute évidence, les éléments prosodiques n'agissent pas tous au même niveau.

(3) *LES DIFFÉRENCES PROSODIQUES ENTRE LE DISCOURS DIRECT SANS VERBE DIRE ET LE DISCOURS DIRECT AVEC VERBE DIRE*

Si on oppose les deux sous-catégories de /discours direct/, le discours direct introduit par le verbe *dire* (DD) et celui introduit sans verbe *dire* (DD∅), la relation prosodie/syntaxe est relative.

D'une part, la pause silencieuse et le débit permettent de distinguer discours direct sans verbe *dire* et discours direct avec verbe *dire*. En effet, la fréquence d'occurrence des pauses silencieuses (et des marqueurs d'attaque) est plus élevée lorsque le verbe *dire* est manquant. De même, le débit est sensible à l'absence du verbe *dire* (DD∅). La vitesse d'articulation est plus lente en discours citant (VA1) qu'en discours cité (VA2) alors que c'est l'inverse lorsque le discours rapporté, directement ou indirectement, est introduit par le verbe *dire*. La vitesse d'articulation médiane de l'énoncé (VAÉ) est également plus lente qu'en DD et qu'en DI.

D'autre part, la mélodie n'est pas significativement différente dans ces deux sous-catégories de discours direct. Ainsi, on ne peut lier les fluctuations de F_0 observées ci-haut (par exemple le fait que la transition soit aussi souvent plate que montante) à la présence ou à l'absence du verbe *dire*. De plus, si globalement, la présence d'une pause silencieuse favorise (i) la diminution des transitions plates en DD et l'augmentation des transitions montantes en DDø; (ii) la diminution des profils montants dans les deux sous-catégories de discours direct, la différence des écarts médians entre les énoncés sans pause silencieuse et ceux avec pause silencieuse n'est pas significative.

Ces observations quant à la relation entre la prosodie et la syntaxe confirment l'aspect relatif du lien qui unit ces deux niveaux linguistiques. D'une part, la stabilité syntaxique du discours indirect se reflète, du moins partiellement, dans une certaine stabilité prosodique. D'autre part, l'instabilité syntaxique du /discours direct/ n'explique pas toutes les variations prosodiques. Même si la fréquence de la pause silencieuse à la transition ainsi que le ralentissement du débit sont fonction de l'absence du verbe *dire*, dans l'ensemble, la mélodie du discours direct avec verbe *dire* n'est pas significativement distincte de celle du discours direct sans verbe *dire*. Cet état de fait m'a amenée à chercher les composantes de la «richesse», ou de la variabilité, fréquencielle du /discours direct/ en conversation courante dans une valeur discursive, la valeur narrative de l'énoncé rapporté.

(4) LES DIFFÉRENCES PROSODIQUES ENTRE LES ÉNONCÉS À VALEUR NARRATIVE FORTE ET LES ÉNONCÉS À VALEUR NARRATIVE FAIBLE

La valeur discursive a été explorée à travers la valeur narrative, à partir de la typologie de VINCENT et DUBOIS (1996), en raison des possibilités systématiques d'attribution d'une valeur narrative à un énoncé.

La mélodie du discours indirect est relativement stable, elle se conforme le plus souvent à la mélodie de l'énoncé déclaratif. Or la mélodie du discours indirect n'est pas véritablement influencée par la valeur narrative La différence de débit entre le discours citant (VA1) et le discours cité (VA2) des discours indirects ne semble pas non plus très liée au type d'emploi discursif, bien qu'une légère tendance à être plus rapide en discours citant ait été décelée dans le type d'emploi discursif de la *reproduction*. Parallèlement, le pourcentage de réussites dans la perception des énoncés en discours indirect ne varie pas en fonction du type d'emploi discursif.

En /discours direct/, la variation fréquencielle observée à la transition est fonction de la valeur narrative. Ainsi, les énoncés dont la valeur narrative est la plus forte présentent plus de transitions montantes (globalement, 5 énoncés sur 10 ont une transition montante lorsque la valeur narrative est forte par rapport à 1 énoncé sur 10 lorsque la valeur narrative est faible). Le débit de l'énoncé (VAÉ) est également plus rapide dans le type d'emploi discursif de la *reproduction*. En revanche, la différence de débit entre le discours citant (VA1) et le discours cité (VA2) des /discours directs/ n'est pas significative. Cette corrélation entre les caractéristiques prosodiques et la valeur narrative se fait finalement sentir dans le pourcentage de réussites dans la perception du /discours direct/ : il passe de 75% en *reproduction* à 58% en *assertion*.

Ce type de corrélation entre prosodique et discursif n'a, à ma connaissance, jamais été établi auparavant. Certes la valeur narrative ici utilisée ne couvre pas entièrement les valeurs discursives attribuables au discours rapporté. Il n'en demeure pas moins que les résultats obtenus démontrent la nécessité d'études acoustiques fines en même temps que de classifications discursives systématiques. Ils ouvrent une piste de recherche des plus intéressantes pour les études prosodiques en général. Selon moi, il s'agit là d'un point important dans cette recherche.

(5) *LES DIFFÉRENCES PROSODIQUES ET LES FACTEURS SOCIAUX*

Dans l'ensemble, il y a peu de corrélation entre les caractéristiques sociales des locuteurs (sexe, âge et statut professionnel) et la prosodie distinctive des deux types de discours rapporté. En /discours direct/, la prosodie dépend beaucoup plus d'un élément discursif, la valeur narrative des paroles rapportées, tel que mentionné dans la réponse à la question (4). En ce qui concerne le discours indirect, la chute à la transition demeure d'abord et avant tout liée à la structure syntaxique de la forme indirecte, structure en tout point semblable à l'énoncé déclaratif. Cette chute semble cependant s'amplifier chez les hommes plus âgés. Il faudrait toutefois des vérifications plus serrées avant d'affirmer que la corrélation est significative. Et si tel était le cas, quelle en serait l'explication ?

Observées dans plusieurs langues et dans plusieurs recherches, les corrélations entre facteurs prosodiques et facteurs sociaux sont bien réelles, mais la nature fluctuante de la prosodie du spontané, particulièrement de l'intonation, ainsi que la variété des méthodes utilisées exigent de la circonspection dans les comparaisons d'une étude à l'autre. De plus, l'interprétation des résultats peut dangereusement glisser vers la subjectivité. À mon avis, que l'étude prosodique soit faite dans une perspective théorique ou appliquée, il importe de vérifier dans quelle mesure la prosodie est soumise à des facteurs sociaux. Quant aux explications des corrélations observées, il est souvent difficile de les fonder sur des faits objectifs.

(6) *LA PERCEPTION DES DIFFÉRENCES PROSODIQUES*

Globalement, il apparaît que lorsque des caractéristiques prosodiques tendancielles du discours rapporté sont présentes (ici F_0 et pause silencieuse), elles favorisent la reconnaissance du type de discours rapporté. Les marques prosodiques du discours indirect sont assez caractéristiques : lorsque l'information prosodique originale est

donnée, le taux de perception est relativement stable alors que lorsque l'information prosodique est manipulée en introduisant des caractéristiques du /discours direct/, le taux de perception diminue sensiblement.

Les marques prosodiques du /discours direct/ sont plus fluctuantes : lorsque l'information prosodique originale est donnée, le taux de perception est variable et lorsque l'information prosodique est manipulée en renforçant les caractéristiques prosodiques, le taux de perception ne s'en trouve généralement pas sensiblement affecté. Ou bien les manipulations ont été maladroites, ou les domaines ont été mal identifiés, ou la variété des motifs discursifs du /discours direct/ demande également une variété de moyens plus étendue que les seuls facteurs prosodiques de la fréquence et de la pause silencieuse; l'intensité joue fort probablement un rôle important. De plus, des moyens paralinguistiques comme le rire, l'hésitation et la qualité de la voix sont certainement mis à contribution, de même que le système mimo-gestuel.

En /discours direct/, les énoncés sont reconnus comme porteurs d'une attitude d'*accord* ou de *désaccord*. Bien sûr, le pourcentage de reconnaissance est supérieur lorsque l'information est intégrale, mais la classification dans une attitude est maintenue lorsque seule l'information prosodique est donnée. Cependant, dans les 6 énoncés observés, F_0 et débit ne jouent pas de rôle significatif dans la reconnaissance de l'attitude. D'autres facteurs prosodiques, telle l'intensité, ou paralinguistiques, telle la qualité de la voix, sont probablement mis à contribution. En discours rapporté, la valeur discursive d'attitude s'ajoute à celle de la valeur narrative et mériterait une investigation approfondie.

Étant donné les résultats obtenus dans l'analyse de la production, j'estime maintenant qu'il a été un peu téméraire de croire que les seules modifications fréquencielles et pausales pourraient améliorer considérablement le taux de perception des /discours directs/

jugés atypiques. Mais les expériences de perception effectuées ont au moins la qualité de le démontrer.

L'étude de la prosodie du discours rapporté spontané a permis d'observer des relations tendancielles entre les divers niveaux linguistiques, mais jamais des relations absolues. Il apparaît que la prosodie n'est captive ni de la syntaxe, ni de la sémantique, ni d'un facteur social pas plus d'ailleurs que d'une attitude... elle s'associe tantôt à l'un, tantôt à l'autre, tantôt à plusieurs.

L'absolu séduit, mais il ne résiste à aucune description approfondie de la parole de tous les jours.

BIBLIOGRAPHIE

ABADJIEVA, E., I.R. MURRAY et J.L. ARNOTT (1993): «Applying analysis of human emotional speech to enhance synthetic speech», *Eurospeech 93 ESCA Proceedings*, Berlin, pp. 909-911.

ALLEN, G.D. (1972): «The location of rhythmic stress beats in English: an experimental study», *Language and Speech*, vol. 15, n° 1, pp. 72-100.

ARRIVÉ, M., F. GADET et M. GALMICHE (1986): *La grammaire d'aujourd'hui*, Paris, Flammarion, 719 p.

AUTHIER, J. (1978): «Les formes du discours rapporté», *DRLAV*, vol. 17, pp. 1-87.

BALIGAND, R. et E. JAMES (1973): «Les structures mélodiques de la phrase interrogative lexicale en franco-ontarien», in A. Grundstrom et P. R. Léon (dir.), *Interrogation et intonation*, Studia Phonetica, vol. 8, Montréal, Didier, pp. 123-145.

BARRY, W.J. (1981): «Prosodic functions revisited again !», *Phonetica*, vol. 38, n^{os} 5-6, pp. 320-340.

BISSONNETTE, S. (1995): *Étude du registre en fonction de la variété de français (Québec versus France)*, mémoire de premier cycle, Université du Québec à Chicoutimi, 53 p.

BLAUUW, E. (1995): «On the perceptual classification of spontaneous and read speech», *Actes du XIII^ème Congrès des Sciences phonétiques*, vol. 3, Stockholm, pp. 254-257.

BOLINGER, D.L. (1958): «Intonation and grammar», *Language Learning*, vol. 8, pp. 31-37.

BOLINGER, D.L. (1964): «Intonation as a universal», in H.G. Lunt(dir.), *Proceedings of the Ninth International Congress of Linguists*, Cambridge (Mass.), The Hague, Mouton, pp. 833-848

BOUDREAULT, M. (1968): *Rythme et mélodie de la phrase parlée en France et au Québec*, Paris, Klincksieck et Québec, Presses de l'Université Laval, 273 p.

BOURDON, B. (1892): *L'expression des émotions et des tendances dans le langage*, thèse de doctorat, Paris, Félix Alcan éditeur.

BRAZIL, D., M. COULTHARD et C. JOHNS (1980): *Discourse Intonation and Language Teaching*, London, Longman.

BREND, R.M. (1972): «Male-female intonation patterns in American English», *Proceedings of the Seventh International Congress of Phonetic Sciences 1971*, The Hague, Mouton. Réimprimé in B. Thorne et N. Henley (dir.), 1975, *Language and Sex: Difference and Domination*, Rowley (Mass.), Newbury House Publishers Inc., pp.84-87.

BRUCE, G. et P. TOUATI (1990): «On the analysis of prosody in spontaneous dialogue», *Working Papers*, n° 36, Lund, Dept. of Linguistics, pp. 37-55.

BYRD, D. (1994): «Relations of sex and dialect to reduction», *Speech Communication*, vol. 15, n^os 1-2, pp. 39-54.

CHARAUDEAU, P. (1992): *Grammaire du français classique et moderne*, Paris, Hachette, 927 p.

CHUNG, S.J. (1995): «An acoustic and perceptual study on the emotive speech in Korean and French», *Actes du XIII^ème Congrès des Sciences phonétiques*, vol. 1, Stockholm, pp. 266-269.

COHEN, M. (1989): *Phonological structures for speech recognition*, Ph. D. dissertation, Berkeley, University of California.

COUPER-KUHLEN, E. (1986): *An Intoduction to English Prosody*, Tübingen, Niemeyer, 239 p.

CRONBACH, L.J. (1951): «Coefficient Alpha and the internal structure of tests», *Psychometrika*, vol. 16, pp.297-334.

CRUTTENDEN, A. (1970): «On the so-called grammatical function of intonation», *Phonetica*, vol. 21, n° 3, pp. 182-192.

CRUTTENDEN, A. (1986): *Intonation*, Cambridge, Cambridge University Press, 214 p.

CRYSTAL, D. (1969): *Prosodic Systems and Intonation in English*, Cambridge, University Press of Cambridge, 381 p.

CRYSTAL, D. (1975): *The English tone of voice. Essays in intonation, prosody and paralanguage*, Londres, Edward Arnold, 189 p.

CRYSTAL, D. (1991): *A Dictionary of Linguistics and Phonetics*, 3^e édition revue et augmentée, Oxford, Basil Blackwell, 389 p.

DELL, F. (1984): «L'accentuation dans les phrases en français» in F. Dell, D. Hirst et J.-R. Vergnaud (dir.), *Forme sonore du langage : structure des représentations en phonologie*, Paris, Hermann, pp. 65-122.

DEMERS, M. (1996): *Socioprosodie, syntaxe et discours. Le cas du discours rapporté en français québécois oral spontané*, thèse de doctorat, Québec, Université Laval, 333 p.

DESHAIES, D. et Cl. PARADIS (1995): «About variable thresholds accounting for the perception of prominent syllables in spontaneous speech», communication présentée à NWAVE, University of Pennsylvania.

DI CRISTO, A. (1985): *De la microprosodie à l'intonosyntaxe*, thèse de doctorat soutenue en 1978, Aix-en-Provence, Publications de l'Université de Provence, 2 tomes, 841 p.

DI CRISTO, A. (à paraître): «Intonation in French», in D.J. Hirst et A. Di Cristo (dir.), *Intonation Systems*, Cambridge, Cambridge University Press.

FÓNAGY, I. (1987): «Semantic diversity in intonation», *Actes du XI^ème Congrès des Sciences phonétiques*, vol. 2, Tallinn, URSS, pp. 468-471.

FORGET, D. (1980): *Analyse sémantique et pragmatique du discours rapporté*, thèse de doctorat inédite, Université de Montréal, 317 p.

FOX, A. (1973): «Tone-Sequences in English», *Archivum Linguisticum*, vol. 4 (New series), pp.17-26.

FRANÇOIS, F. (1966): «La description linguistique», in A. Martinet (dir.), *Le langage*, Encyclopédie de la Pléiade, Paris, Gallimard, pp. 171-281.

GAULMYN, M.-M. de (1992): «Grammaire du français parlé. Quelques remarques autour du discours rapporté», in A.-M. Joussaud et J. Pétrissans (dir.), Actes du Congrès de l'ANEFLE, *Grammaire et français langue étrangère*, Grenoble, ANEFLE, pp.22-33.

GENDRON, J.-D. (1966): *Tendances phonétiques du français parlé au Canada*, Paris, Klincksieck et Québec, Presses de l'Université Laval, 254 p.

GOFFMAN, E. (1964): «The neglected situation», in J.J. Gumperz et D. Hymes (dir.), *American Anthropologist*, publication spéciale, vol. 66, n° 6, partie 2, pp. 133-137.

GOLDMAN-EISLER, F. (1968): *Psycholinguistics : Experiments in spontaneous Speech*, Londres, Academic Press, 169 p.

GREVISSE, M. (1986): *Le bon usage*, 12 ième édition refondue par A. Goosse, Paris, Éditions Duculot, 1768 p.

GROSJEAN, F. et A. DESCHAMPS (1972): «Analyse des variables temporelles du français spontané», *Phonetica*, vol. 26, n° 3, pp. 129-156.

GROSJEAN, F. et A. DESCHAMPS (1975): «Analyse contrastive des variables temporelles de l'anglais et du français : vitesse de parole et variables composantes, phénomènes d'hésitation», *Phonetica*, vol. 31, n°s 3-4, pp. 144-184.

GUAÏTELLA, I. (1991): *Rythme et parole : comparaison critique du rythme de la lecture oralisée et de la parole spontanée*, thèse de doctorat inédite, Aix-en-Provence, Université d'Aix-en-Provence, 420 p.

GUAITELLA, I. (1992): «Hésitations vocales en parole spontanée : réalisations acoustiques et fonctions rythmiques», *Travaux de l'Institut de Phonétique d'Aix*, n° 14, pp. 113-130.

GUMPERZ, J.J. et D. HYMES (dir). (1964): «The ethnology of communication, *American Anthropologist*, publication spéciale, vol. 66, n° 6, partie 2, pt.ii.

GUY, G., B. HORVATH, J. VONWILLER, E. DAISLY et I. ROGERS (1986): «An intonational change in progress in Australian English», *Language and Society*, vol. 15, n° 1, pp. 23-52.

GÅRDING, E. (1977): «The importance of turning points for the pitch patterns of Swedish accents», in L.M. Hyman (dir.), *Studies in stress and accent*, Los Angeles, Southern California Occasional Papers in Linguistics, n° 4, pp. 27-35.

HALLIDAY, M.A.K. (1967): *Intonation and Grammar in British English*, The Hague, Mouton.

HALLIDAY, M.A.K. (1970): «Language structure and language function», in J. Lyons (dir.), *New Horizons in Linguistics*, Harmondsworth, Penguin, pp.140-165.

HALLIDAY, M.A.K. (1976): «Functions and universals of language», in G.R. Kress (dir.), *Halliday: System and Function in Language*, London, Oxford University Press, pp.26-31.

HART, J. 't (1981): «Differential sensivity to pitch distance, particularly in speech», *The Journal of the Acoustical Society of America*, vol. 69, n° 3, pp. 811-821.

HART, J. 't, R. COLLIER et A. COHEN (1990): *A perceptual study of intonation : an experimental-phonetic approach to speech melody*, Cambridge, Cambridge University Press, 212 p.

HIRST, D.J. et A. DI CRISTO (à paraître): «A survey of intonation systems», in D.J. Hirst et A. Di Cristo (dir.), *Intonation Systems*, Cambridge, Cambridge University Press.

HOLDER, M. (1968): «Étude sur l'intonation comparée de la phrase énonciative en français canadien et en français standard», in P.R. Léon (dir.), *Recherches sur la structure phonique du français canadien*, Studia Phonetica, vol. 1, Montréal, Didier, pp. 175-191.

HOLLIEN, H. et T. SHIPP (1972): «Speaking fundamental frequency and chronological age in males», *Journal Speech Hearing Research*, vol. 15, pp. 155-159.

KEY, M.R. (1972): «Linguistic behavior of male and female», *Linguistics*, vol. 88, pp. 15-31.

KONOPCZYNSKI, G. (1986): *Du prélangage au langage : acquisition de la structuration prosodique*, thèse de doctorat d'état, Université de Strasbourg.

KONOPCZYNSKI, G. (1990) : «Le langage émergent : comment la mélodie devient intonation entre 8 et 24 mois», in N. Zavialoff (dir.), *La lecture : de la neurobiologie à la pédagogie*, Tome 1, Paris, L'Harmattan, pp. 329-341.

LABOV, W. (1964): «Phonological Correlates of Social Stratification» in J.J. Gumperz et D. Hymes (dir.), *American Anthropologist*, publication spéciale, vol. 66, n° 6, partie 2, pp. 164-176.

LADD, D.R. Jr. (1980): *The Structure of Intonational Meaning. Evidence from English*, Bloomington, Indiana University Press.

200

LADD, D.R. Jr., K.R. SCHERER et K. SILVERMAN (1984): «An integrated approach to studying intonation and attitude», in C. Johns-Lewis (dir.), *Studies in Intonation and Discourse*, London, Croom Helm.

LAFOREST, M. (1995): *Autour de la narration*, Québec, Nuit blanche éditeur, 172 p.

LAKOFF, R. (1975): *Language and Woman's Place*, New York, Harper and Row, 83 p.

LEANDRI, S. (1993): «Prosodic aspects of reported speech», in D. House et P. Touati (dir.) *ESCA Workshop on Prosody 1993*, n° 41, Lund (Suède), Dept of Linguistics and Phonetics, Lund University, pp. 152-155.

LE GOFFIC, P. (1993): *Grammaire de la phrase française*, Paris, Hatier, 591 p.

LEON, P.R. et Ph. MARTIN (1970): *Prolégomènes à l'étude des structures intonatives*, Studia Phonetica, vol. 2, Montréal, Didier, 225 p.

LIEB, H.-H. (1980): «Intonation als Mittel verbaler Kommunikation», *Linguistische Berichte*, n° 68, pp.34-48.

LIEBERMAN, P. (1965): «On the acoustic basic of the perception of intonation by linguists», *Word*, vol. 21, pp. 40-54.

LIEBERMAN, P., W. KATZ, A. JONGMAN, R. ZIMMERMAN et M. MILLER (1985): «Measures of the sentence intonation of read and spontaneous speech in American English», *The Journal of the Acoustical Society of America*, vol. 77, n° 2, pp.649-657.

201

LIEBERMAN, P. et S.D. MICHAELS (1962): «Some aspects of fundamental frequency, envelope amplitude and the emotional content of speech», *The Journal of the Acoustical Society of America*, vol. 34, n° 7, pp. 922-927.

MAINGUENEAU, D. (1981): *Approche de l'énonciation en linguistique française*, Paris, Hachette, 127 p.

MAINGUENEAU, D. (1986): *Éléments de linguistique pour le texte littéraire*, Paris, Bordas, 156 p.

MAINGUENEAU, D. (1987): *Nouvelles tendances en analyse du discours*, Paris, Hachette, 144 p.

MALECOT, A., R. JOHNSTON et P.A. KIZZIAR (1972): « Syllabic rate and utterance length in French», *Phonetica*, vol. 26, n° 4, pp.235-251.

MAROUZEAU, J. (1924): «Accent affectif et accent intellectuel» *Bulletin de la Société linguistique de Paris*, vol. 25, pp. 80-86.

MARTIN, Ph. (1981): «Pour une théorie de l'intonation. L'intonation est-elle une structure congruente à la syntaxe ?» *L'intonation, de l'acoustique à la sémantique*, Paris, Klincksieck, pp. 234-271.

MAURY, N. et P. WRENN (1973): «L'interrogation mélodique en français canadien de l'Ontario», in A. Grundstrom et P.R. Léon (dir.), *Interrogation et intonation en français standard et en français canadien*, Studia Phonetica, vol. 8, Montréal, Didier, pp.99-121.

MCCONNELL-GINET, S. (1978): «Intonation in a Man's World», *Journal of Women in Culture and Society*, vol. 3, n° 3, pp. 541-559.

McDOWALL, J.J. (1974): «The reliability of ratings by linguistically untrained subjects in response to stress in speech», *Journal of Psycholinguistic Research*, vol. 3, pp. 247-259.

MEHTA, G. et A. CUTLER (1988): «Detection of target phonemes in spontaneous and read speech», *Language and Speech*, vol. 31, n° 2, pp. 135-156.

MERTENS, P. (1987): *L'intonation du français. De la description linguistique à la reconnaissance automatique*, thèse de doctorat inédite, Louvain, Katholieke Universiteit Leuven, 323 p.

MERTENS, P. (1990): «Intonation» in C. Blanche-Benveniste, M. Bilger, C. Rouget et K. van den Eynde (dir.), *Le français parlé*, Paris les Éditions du CNRS, pp. 159-175.

MERTENS, P. (1991): «Local prominence of acoustic and psychoacoustic functions and perceived stress in French», *Actes du XII^e Congrès international des sciences phonétiques*, vol. 3, Aix-en-Provence, pp. 218-221.

MERTENS, P. (1992): «L'accentuation des syllabes contiguës», *ITL Review of Applied Linguistics*, n^os 95-96, Louvain, pp. 145-165.

MERTENS, P. (1993): «Accentuation, intonation et morphosyntaxe», *Travaux de Linguistique*, vol. 26, pp. 21-67.

MILLER, J.L., F. GROSJEAN et C. LOMANTO (1984): «Articulation rate and its variability in spontaneous speech : A reanalysis and some implications», *Phonetica*, vol. 41, n° 4, pp. 215-225.

MORA, E. (1990): «Phonostylistique de l'intonation : différenciations dues au milieu social et au sexe des locuteurs», *Revue québécoise de linguistique*, vol. 19, n° 2, pp.73-91.

MOREL, M.A. (1996): «Le discours rapporté direct dans l'oral spontané», *Cahiers du français contemporain*, n° 3, pp. 77-90.

MOZZICONACCI, S. (1995): «Pitch variations and emotions in speech», *Actes du XIII^ème^ Congrès des Sciences phonétiques*, vol. 1, Stockholm, pp. 178-181.

MYSAK, E. (1959): «Pitch and duration characteristics of older males», Journal Speech Hearing *Reasearch*, vol. 6, pp. 164-170.

O'CONNOR, J.D. et G.F. ARNOLD (1973): *Intonation of Colloquial English*, London, Longman.

OUELLON, C., Cl. PARADIS et L. DUCHESNE (1991): «Les marqueurs acoustiques de l'énoncé en français québécois.», *Actes du XII^e^ Congrès international des sciences phonétiques*, vol. 2, Aix-en-Provence, pp. 430-433.

PARADIS, Cl. et D. DESHAIES (1990): «Rules of stress assignment in Québec French : Evidence from perceptual data» *Language Variation and Change*, vol. 2, Cambridge, Cambridge University Press, pp. 135-154.

PASDELOUP, V. (1988): «Essai d'analyse du système accentuel en français : distribution de l'accent secondaire», *Actes des XVII^èmes^ Journées d'Études sur la Parole*, Nancy.

PASDELOUP, V. (1990): *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*, thèse de doctorat inédite, Aix-en-Provence, Université de Provence I.

PIERREHUMBERT, J. (1980): *The phonology and phonetics of English intonation*, thèse de doctorat inédite, Cambridge, Massachusetts Institute of Technology, 401 p.

RAMIG, L. (1983): «Effects of physiological aging on speaking and reading rates», *Journal of Communication Disorders*, vol. 16, pp. 217-226.

RIE, J. van et R. van BEZOOIJEN (1995): «Perceptual characteristics of voice quality in dutch males and females from 9 to 85 years», *Actes du XIII[ème] Congrès des Sciences phonétiques*, vol. 2, Stockholm, pp.290-293.

ROSSI, M. (1971): «Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole», *Phonetica*, vol. 23, n° 1, pp. 1-33.

ROSSI, M. (1978): «La perception des glissandos descendants dans les contours prosodiques», *Phonetica*, vol. 35, n° 1, pp. 11-40.

ROSSI, M. (1979): «Le français, langue sans accent ?», in I. Fónagy et P. R. Léon (dir.), *L'accent en français contemporain*, Studia Phonetica, vol. 15, Montréal, Didier, pp. 13-52.

ROSSI, M. (1985): «L'intonation et l'organisation de l'énoncé», *Phonetica*, vol. 42, n[os] 2-3, pp.135-153.

RYALLS, J., G. LE DORZE, N. LEVER, L. OUELLET et C. LARFEUIL (1994): «The effects of age and sex on speech intonation and duration for matched statements and questions in French», *The Journal of the Acoustical Society of America*, Lettres à l'Éditeur, vol. 95, n° 4, pp. 2274-2276.

RYAN, W.J. (1972): «Acoustic aspects of the aging voice», *Journal of Gerontology*, vol. 27, n° 2, pp. 265-272.

SACHS, J. (1975): «Cues to the identification of sex in children's speech», in B. Thorne et N. Henley, *Language and Sex : Difference and Dominance*, Rowley (Mass.), Newbury House Publishers Inc., pp. 152-171.

SANKOFF, D., G. SANKOFF, S. LABERGE et M. TOPHAM (1976): «Méthodes d'échantillonnage et utilisation de l'ordinateur dans l'étude de la variation linguistique», *Cahiers de linguistique de l'Université du Québec*, n° 6, pp. 85-125.

SANTERRE, L. (1990): «La condition de non-contiguïté accentuelle en français : théorie et pratique», *Revue québécoise de linguistique*, vol. 19, n° 2, Université du Québec à Montréal, pp. 39-57.

SILVERMAN, K., K.R. SCHERER et D.R. Jr. LADD (1983): «Vocal cues to attitude in non-emotional speech», Ms, Dept. of Psychology, University of Giessen (Allemagne).

SMITH, P.M. (1979): «Sex markers in speech», in K.R. Scherer et H. Giles (dir.), *Social Markers in Speech*, Cambridge, Cambridge University Press, pp. 109-146.

TENCH, P. (1990): *The roles of intonation in English discourse*, Frankfurt am Main, Peter Lang, 534 p.

THIBAULT, L. (1994): *Étude exploratoire du rythme en français québécois*, mémoire de maîtrise inédit, Québec, Université Laval, 196 p.

THIBAULT, P. et D. VINCENT (1990): *Un corpus de français parlé : Montréal 1984*, Québec, Presses de l'Université Laval, 145 p.

THORNE, B. (1983): «Language, gender and society : opening a second decade of research», in B. Thorne, C. Kramarae et N. Henley (dir.), *Language, gender and society*, Rowley (Mass.), Newbury House Publishes Inc., pp. 7-24.

TOUATI, P. (1987): *Structures prosodiques du suédois et du français. Profils temporels et configurations tonales*, Travaux de l'Institut de linguistique de Lund XXI, Lund, Lund University Press, 208 p.

ULDALL, E. (1960): «Attitudinal meanings conveyed by intonation contours», *Language and Speech*, vol. 3, pp. 223-234.

ULDALL, E. (1964): «Dimensions of meaning in intonation», in D. Bolinger (dir.), *Intonation. Selected readings*, Harmondsworth, Penguin, pp.250-259.

VICHER, A. et D. SANKOFF (1989): «The emergent syntax of pre-sentential turn-openings», *Journal of Pragmatics*, vol. 13, pp.91-97.

VINAY, J.-P. (1955): «Aperçu des études de phonétique canadienne», *Études sur le parler français au Canada*, Québec, Les Presses de l'Université Laval, pp.61-82.

VINCENT, D. et S. DUBOIS (1996): *Le discours rapporté*, Québec, Nuit blanche éditeur, 123 p.

VROOMEN, J., R. COLLIER et S. MOZZICONACCI (1993): «Duration and intonation in emotional speech», *Eurospeech '93 ESCA Proceedings*, Berlin, pp. 577-580.

WAGNER, R.-L. et J. PINCHON (1962): *Grammaire du français classique et moderne*, Paris, Hachette, 640 p.

YULE, G. (1980): «Speakers' topics and major paratones», *Lingua*, vol. 52, pp. 33-47.

ANNEXES

***ANNEXE A* : Expérience de perception 1**

***ANNEXE B* : Expérience de perception 9**

ANNEXE A **EXPÉRIENCE de PERCEPTION 1**
Différences grammaticales

EXPLICATION

Pour citer les paroles de quelqu'un, on peut s'y prendre de deux façons: en utilisant exactement les mêmes mots, c'est ce qu'on appelle du **discours direct (DIR)**; ou bien en modifiant quelque peu ce qui a été dit, c'est ce qu'on appelle alors du **discours indirect (INDIR)**.

PAR EXEMPLE :
Discours direct (DIR)
Il m'a dit: «Les tests de perception sont utiles.»
Il s'agit donc de paroles citées telles quelles.
Parfois, le verbe *dire* est absent.
Puis très compréhensif : «J'accepte d'en passer.»
Discours indirect (INDIR)
*Il m'a dit **que** les tests de perception **étaient** utiles.*
Dans ce cas, il s'agit de paroles rapportées avec des modifications.

CONSIGNES

1. Vous allez entendre des phrases dites par des personnes différentes (4 ou 5). C'est de l'oral spontané, donc sorti du contexte et pas toujours clair. Il faut parfois accepter de répondre intuitivement.
2. Chaque phrase est répétée deux (2) fois d'affilée; il y a une seconde d'intervalle entre les répétitions, trois (3) secondes entre les phrases et 6 secondes entre chaque groupe de cinq (5) phrases.
3. On vous demande d'identifier la phrase entendue comme étant du **discours direct (DIR)** ou du **discours indirect (INDIR).**

   Par exemple, pour le numéro 1:

   - si vous croyez que la phrase 1 correspond à du **discours direct (DIR)**, cochez (√):
     1. DIR ☑ INDIR ☐
   - Si vous croyez plutôt que la phrase 1 correspond à du **discours indirect (INDIR)**, cochez (√):
     1. DIR ☐ INDIR ☑
4. À titre d'exemple, voici la phrase qui correspond au numéro 1 de la feuille-réponse.
5. Est-ce qu'il y a des questions ?

FEUILLE-RÉPONSE

Cocher (√) le type de discours correspondant à la phrase perçue.

EXEMPLE					
1.	DIR ☐	INDIR ☐	16.	DIR ☐	INDIR ☐
DÉBUT DU TEST					
2.	DIR ☐	INDIR ☐	17.	DIR ☐	INDIR ☐
3.	DIR ☐	INDIR ☐	18.	DIR ☐	INDIR ☐
4.	DIR ☐	INDIR ☐	19.	DIR ☐	INDIR ☐
5.	DIR ☐	INDIR ☐	20.	DIR ☐	INDIR ☐
6.	DIR ☐	INDIR ☐	21.	DIR ☐	INDIR ☐
7.	DIR ☐	INDIR ☐	22.	DIR ☐	INDIR ☐
8.	DIR ☐	INDIR ☐	23.	DIR ☐	INDIR ☐
9.	DIR ☐	INDIR ☐	24.	DIR ☐	INDIR ☐
10.	DIR ☐	INDIR ☐	25.	DIR ☐	INDIR ☐
11.	DIR ☐	INDIR ☐	26.	DIR ☐	INDIR ☐
12.	DIR ☐	INDIR ☐	27.	DIR ☐	INDIR ☐
13.	DIR ☐	INDIR ☐	28.	DIR ☐	INDIR ☐
14.	DIR ☐	INDIR ☐	29.	DIR ☐	INDIR ☐
15.	DIR ☐	INDIR ☐	30.	DIR ☐	INDIR ☐

ANNEXE B **EXPÉRIENCE de PERCEPTION 9**
Différences «attitudinales»

EXPLICATION

Quand on rapporte les paroles de quelqu'un, on peut être en ACCORD ou en DÉSACCORD avec ce que cette personne a dit. Le «ton» utilisé par le rapporteur peut souvent être un indice de son accord ou de son désaccord.

PAR EXEMPLE:
Il m'a dit: «Les tests de perception sont utiles.»
(ACCORD: Je les trouve utiles aussi.)
(DÉSACCORD: Moi je les trouve inutiles.)

CONSIGNES

1. Vous allez entendre six (6) phrases.
2. Chaque phrase est répétée trois (3) fois d'affilée; il y a une pause d'une seconde entre chaque répétition et de cinq (5) secondes entre chaque phrase.
3. Chacune des phrases est transcrite sur la feuille-réponse; sous chacune des transcriptions, deux (2) interprétations vous sont proposées **(en gras, entre parenthèses)**: la première explicite un ACCORD de celui qui parle avec les paroles qu'il cite; la seconde, son DÉSACCORD.
   Par exemple:
   1. C'est une fille qui était pas rodée pi tout 'suite (pause): «Ces deux-là dehors.»
   - □ **(ACCORD: Il fait bien de les mettre à la porte.)**
   - □ **(DÉSACCORD: Moi je leur aurais donné une chance.)**
4. On vous demande:
   - de lire les deux interprétations proposées tout en écoutant la mélodie;
   - de cocher (√) l'interprétation, ACCORD ou DÉSACCORD, correspondant à la phrase entendue.
5. À titre d'exemple, vous allez entendre la phrase qui correspond au no 1 de la feuille-réponse.
6. Est-ce qu'il y a des questions?
7. Début du test.

FEUILLE-RÉPONSE

Cocher (√) l'interprétation perçue.

EXEMPLE

1. C'est une fille qui était pas rodée pi tout 'suite (pause): «Ces deux-là dehors.»
   - ☐ **(ACCORD: Il fait bien de les mettre à la porte.)**
   - ☐ **(DÉSACCORD: Moi je leur aurais donné une chance.)**

DÉBUT DU TEST

2. Elle lui a dit: «Oui t'as été élevé comme un vrai fou.»
   - ☐ **(ACCORD: C'est vrai qu'il a été mal élevé.)**
   - ☐ **(DÉSACCORD: Moi je trouve qu'il a été bien élevé.)**

3. Les autres m'encourageaient: «Cours cours Bina.»
   - ☐ **(ACCORD: J'étais contente de leurs encouragements.)**
   - ☐ **(DÉSACCORD: Je n'étais pas contente; j'étais épuisée.)**

4. Ils l'mettent sur le tableau: «Pi arrangez-vous avec le tableau.»
   - ☐ **(ACCORD: C'est la bonne façon de faire.)**
   - ☐ **(DÉSACCORD: Ce n'est pas la bonne façon de faire.)**

5. La femme a dit (pause): «Ben on vous donnerait pas ça.»
   - ☐ **(ACCORD: C'est vrai que j'ai l'air jeune.)**
   - ☐ **(DÉSACCORD: Elle a juste voulu me flatter.)**

6. Mon mari il dit: «Ah ben c'est correct.»
   - ☐ **(ACCORD: Il fait bien d'accepter.)**
   - ☐ **(DÉSACCORD: Il devrait refuser.)**

L'étude porte sur les rapports entre prosodie, syntaxe et discours à travers le discours rapporté en français québécois (montréalais) oral spontané. Il s'agit d'une analyse instrumentale de F_0, de la pause et du débit, analyse basée sur 215 énoncés de discours rapporté. À partir d'une classification syntaxique, discours direct et discours indirect, l'analyse de la production démontre une certaine stabilité prosodique du discours indirect et une fluctuation prosodique du discours direct; en revanche, une classification discursive démontre que la fluctuation prosodique du discours direct est souvent fonction de la valeur narrative de l'énoncé. Une analyse de la variation socioprosodique prend en compte le sexe, l'âge et le statut socioprofessionnel; ces facteurs sociaux semblent peu corrélés à la distinction discours direct/discours indirect. L'analyse de la perception montre d'autre part que plus les tendances prosodiques sont fortes, plus le taux de perception du type de discours est élevé. Il ressort que la prosodie n'est captive ni de la syntaxe, ni de la sémantique, ni d'un facteur social. La prosodie est souple et fonctionnelle.